Reto Francioni · Robert A. Schwartz
Editors

Equity Markets in Transition

The Value Chain, Price Discovery, Regulation, and Beyond

Springer

Equity Markets in Transition

Reto Francioni • Robert A. Schwartz
Editors

Equity Markets in Transition

The Value Chain, Price Discovery, Regulation, and Beyond

with the assistance of John Byrne and Stephanie Welkoborsky

Springer

Editors
Reto Francioni
Faculty of Economics and Business Administration
University of Basel
Basel, Switzerland

Robert A. Schwartz
Zicklin School of Business
Baruch College/CUNY
New York, NY, USA

ISBN 978-3-319-45846-5 ISBN 978-3-319-45848-9 (eBook)
DOI 10.1007/978-3-319-45848-9

Library of Congress Control Number: 2017930221

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Abbreviations

ABS	Asset Backed Security
BaFin	Bundesanstalt für Finanzdienstleistungsaufsicht (DE)
Basel III	Third Basel Accord
BCBS	Basel Committee for Banking Supervision
BIS	Bank for International Settlement
CAP	Client asset protection
CCP	Central Counterparty
CDO	Collateralized Debt Obligation
CFTC	US Commodities Futures Trading Commission
CH	Clearing House
CMU	Capital Markets Union
CPMI (CPSS)	Committee on Payments and Market Infrastructures (formerly Committee on Payment and Settlement Systems)
CPU	Central Processing Unit
CRD IV/CRR	Capital Requirements Directive/Regulation
CSD	Central Securities Depository
CSD-R	Central Securities Depositories Regulation
DCO	Derivatives clearing organization
DTCC	Depository Trust and Clearing Corporation
DvP	Delivery versus Payment
EBA	European Banking Authority—Regulation (EU) 1093/2010
ECB	European Central Bank
ESFS	European System of Financial Supervision
EIOPA	European Insurance and Occupational Pensions Authority—Regulation (EU) 1094/2010
EMIR	European Market Infrastructure Regulation—Regulation (EU) No 648/2012
EP	European Parliament
ESAs	European Supervisory Authorities
ESMA	European Securities and Markets Authority—Regulation (EU) No 1095/2010

ESRB	European Systemic Risk Board
EU COM	European Commission
FinFraG	Finanzmarktinfrastrukturgesetz (CH)
FIX	Financial Information eXchange
FMI	Financial Market Infrastructure
FOP	Free-of-payment
FPGA	Field Programmable Gate Arrays
FSB (FSF)	Financial Stability Board (formerly Financial Stability Forum)
G20	Group of 20
GAAP	United States Generally Accepted Accounting Principles
GPS	Global Positioning System
GUI	Graphical User Interface
HFT	High Frequency Trading
IAIS	International Association of Insurance Supervisors
IASB	International Accounting Standards Board
ICE	International Commodity Exchange
ICSD	International Central Securities Depository
IFRS	International Financial Reporting Standards
IMF	International Monetary Fund
IOSCO	International Organization of Securities Commissions
IPO	Initial Public Offering
IRS	Interest rate swaps
ISA	Individual segregated account
ISIN	International Securities Identification Number
ISO	International Organization for Standardization
IT	Information Technology
JRE	Java Runtime Environment
LSOC	Legally separated, operationally comingled
MBS	Mortgage Backed Security
MiFID/MiFIR	Markets in Financial Instruments Directive/Regulation
MTF	Multilateral Trading Facility
NCM	Non-clearing member
NTP	Network Time Protocol
OCC	US Office of the Comptroller of the Currency
OECD	Organisation for Economic Co-operation and Development
OSA	Omnibus segregated account
OTC	Over the Counter
OTF	Organized Trading Facilities
PS	Payment system
PTP	Precision Time Protocol
RC	Registered customer
RDMA	Remote Direct Memory Access
Repo	Repurchase agreement
RM	Regulated Markets
RTGS	Real-time gross settlement

SEC	Securities Exchanges Commission (US)
SEF	Swap Execution Facilities
SGX	Singapore Exchange
SI	Systemic Internalizers
SSBs	Standard-Setting Bodies
SSS	Securities Settlement System
STP	Straight through processing
SWIFT	Society for Worldwide Interbank Financial Telecommunication
T2S	TARGET2-Securities
TCP/IP	Transmission Control Protocol/Internet Protocol
UDP	User Datagram Protocol
VPN	Virtual Private Network
WAN	Wide Area Network
WB	World Bank
WpHG	Wertpapierhandelsgesetz
WTO	World Trade Organization

About the Authors

James H. Freis Jr. is Chief Compliance Officer for the Deutsche Börse Group since joining in 2014. He is responsible for ensuring compliance with regulatory requirements and internal policies within Group entities globally, and interactions with relevant supervisory authorities. He previously served in the United States Department of the Treasury, including as Director (CEO) of the Financial Crimes Enforcement Network (FinCEN), at the Bank for International Settlements in Basel, Switzerland, and the Federal Reserve Bank of New York, as well as the international law firm Cleary Gottlieb Steen & Hamilton.

Cord Gebhardt is a director at Deutsche Börse AG in the cash market division at Deutsche Börse Group. As a management board member of the Deutsche Börse-owned Frankfurt Stock Exchange and head of Department Primary Market, he is responsible for the Deutsche Börse's primary market business. After completing legal studies at Justus Liebig University Giessen and the University of Bonn, Dr. Gebhardt first served as policy officer at the German Federal Foreign Office prior to completing his doctorate at the Max Planck Institute for European Legal History in Frankfurt am Main in 1997. In 1996, Dr. Gebhardt joined the legal department at Deutsche Börse AG; from 2000 to 2006, he headed the Securities Listing Admission Department and the Admission Office of the Frankfurt Stock Exchange. From 2006 to early 2013, he was the head of Section Markets and Regulatory at the legal department at Deutsche Börse AG.

Alexandra Hachmeister is Chief Regulatory Officer of Deutsche Börse AG, which she joined in 1999. Since then she served in various positions and departments including Market Design Functionality, Systems Business Development, and Boards & Committees. In her role as Head of Regulatory Strategy she is responsible for the development, implementation, and communication of the Group's global regulatory strategy and positioning. Alexandra is a member of the Group of Economic Advisors of the European Securities and Markets Authority as well as of the Working Group "Financial Transaction Tax" at the German Exchange & Infrastructure Commission and of the Working Group "Role of Financial Services in Society" of the World Economic Forum. She holds a Ph.D. in Market Microstructure from the EBS University in Wiesbaden (Germany) and lectures on regulatory topics at the University of Basel (Switzerland) and the University of Frankfurt (Germany).

Matthias Kluber is managing director at Deutsche Börse with global responsibility for IT infrastructure and operations. Kluber studied mathematics and information technology at the Friedrich-Alexander-Universität of Erlangen-Nuremberg. After earning a German diploma degree in 1990, he joined IBM and participated in the standardization and consolidation of IBM's data centers across Europe. In 1993, Kluber started as a senior consultant for Andersen Consulting (subsequently renamed Accenture) and later as a project manager. At Accenture, he led project teams in London, Berlin, Frankfurt, and Stuttgart. Kluber joined Deutsche Börse in 1998 and, since then, has served in various management positions. His teams of approximately 450 IT professionals are based in Frankfurt, Luxembourg, Chicago, and Prague.

Thomas Laux is Chief Risk Officer of Eurex Clearing, the leading clearing house in Europe for securities and derivatives transactions. At Eurex Clearing, he is responsible for the design, development, and operation of all risk management models and services. Prior to joining the clearing house in 2006, Laux spent 6 years at Deutsche Börse Systems on company-wide strategic projects. He holds a dual MBA from the Fuqua School of Business, Duke University, and the Goethe Business School, Frankfurt, where he was recognized as a Fuqua & Goethe Scholar. Laux also received a degree in mathematics and computer science from the Technical University of Darmstadt.

Martin Reck is the Managing Director at Deutsche Börse AG responsible for the Cash Market Structure and the electronic trading system Xetra®. Currently some 230 European banks and broker firms from 20 countries use the Xetra system. Over 1000 ETFs, 400 ETCs and ETNs, 10,000 equities from 70 countries, 26,000 corporate and government bonds, and one million structured products are traded on Xetra—more than on any other trading venue. Since 1998, Mr. Reck has held a number of leadership and project management positions with Deutsche Börse Group such as Head of Xetra Market Design and Functionality. In 2003, he became a Managing Director and was responsible for Group Functionality, from 2007 on for Group Strategy. Since 2007, he also is a lecturer at Johann Wolfgang Goethe-University Frankfurt, Faculty of Economics and Business Administration, Chair of Corporate Finance. Martin is a German native. He holds a Diploma in Computer Science from the University of Dortmund. He also holds a Ph.D. in Information Systems of the Institute for Information Systems, University of St. Gallen.

Marc Robert-Nicoud is the Chief Executive Officer of Clearstream Holding, a wholly owned company of Deutsche Börse AG. His areas of direct responsibility are strategy, pricing, and communication. He joined Deutsche Börse Group in 2006 as a legal counsel and has since fulfilled various roles with a focus on Clearstream's post-trade services. Previously, Robert-Nicoud was a legal officer at the Swiss Financial Market Supervisory Authority (FINMA). He started his career as an associate with the international law firm, Cleary Gottlieb Steen & Hamilton. He is a member of the New York State Bar and the Quebec Bar and is a common law (L.L.B.) and civil law (B.C.L.) graduate of McGill University in Montreal, Canada. He also completed the Executive Program of the Swiss Finance Institute.

Matthias Stötzel is Head of Regulatory Legal at Deutsche Börse AG. He joined the Group in 2006 and was an advisor to the cash market business of Deutsche Börse Group until 2012 when he was appointed the head of Clearing Legal. At Clearing Legal, he was responsible for the legal advice to the group's central counterparty, Eurex Clearing AG. In 2013, Dr. Stötzel took the reins of Section Trading Legal that provides legal advice of the trading platforms, including Eurex Exchange, Eurex Repo, Eurex Bonds, and the Frankfurt Stock Exchange, operated by Deutsche Börse Group. Before joining Deutsche Börse, Dr. Matthias Stötzel worked in the law firm Schlawien Naab focusing on civil contract law and for the international law firm Gleiss Lutz, specializing in public law, particularly environmental and planning law as well as energy law.

Jan Strecker is a director at Deutsche Börse AG, heading investor relations. Since joining Deutsche Börse in 2000, he has held various positions within the Group. Before joining the investor relations team in 2005, Strecker worked on group-wide M&A projects and in the cash market business development. He has a bachelor's degree in business administration and has completed the Certified European Financial Analyst (CEFA) program.

About the Editors

Reto Francioni has been a professor of applied capital markets theory at the University of Basel since 2006. Since 2013, he has been a member of the Board of Directors of UBS Group, and since June 2016, he is Chairman of the Supervisory Board of Swiss International Airlines as well as Senior Independent Director of the Coca-Cola HBC AG.

Reto Francioni was CEO of Deutsche Börse AG from 2005 to May 2015. From 2002 to 2005, he was Chairman of the Supervisory Board and President of the SWX Group (Stock Exchange of Switzerland). Reto Francioni was co-CEO and Spokesman for the Board of Directors of Consors AG from 2000 to 2002. Between 1993 and 2000, he held various management positions at Deutsche Börse AG, including that of Deputy CEO. From 1992 to 1993, he served in the corporate finance division of Hoffmann-La Roche as a member of the Directorate. Prior to this, he was on the executive board of Association Tripartite Bourses and worked, from 1985 to 1988, for the former Credit Suisse. Reto Francioni started his career in 1981 in the commerce division of Union Bank of Switzerland. He studied and earned his doctorate in law at the University of Zurich.

Robert A. Schwartz is the Marvin M. Speiser Professor of Finance and University Distinguished Professor in the Zicklin School of Business, Baruch College, CUNY. Before joining the Baruch faculty in 1997, he was Professor of Finance and Economics and Yamaichi Faculty Fellow at New York University's Leonard N. Stern School of Business, where he had been a member of the faculty since 1965. In 1966, Professor Schwartz received his Ph.D. in Economics from Columbia University. His research is in the area of financial economics, with a primary focus on the structure of securities markets. He has published dozens of refereed journal articles, 12 edited books, and 8 authored books, including *Micro Markets: A Market Structure Approach to Microeconomic Analysis* (Wiley & Sons, 2010); *Mastering the Art of Equity Trading Through Simulation: The TraderEx Course* (coauthored with Gregory Sipress and Bruce Weber, Wiley & Sons, 2010); *The Equity Trader Course* (coauthored with Reto Francioni and Bruce Weber, Wiley & Sons, 2006); and *Equity Markets in Action: The Fundamentals of Liquidity, Market Structure and*

Trading (coauthored with Reto Francioni, Wiley & Sons, 2004). He has served as a consultant to various market centers, including the New York Stock Exchange, the American Stock Exchange, Nasdaq, the London Stock Exchange, Instinet, the Arizona Stock Exchange, Deutsche Börse, the Bolsa Mexicana, and the Istanbul Stock Exchange. From April 1983 to April 1988, he was an associate editor of *The Journal of Finance*, and he is currently an associate editor of the *Review of Quantitative Finance and Accounting* and the *Review of Pacific Basin Financial Markets and Policies*, is a member of the advisory boards of *International Finance* and *The Journal of Trading*, and is on the editorial board of *International Journal of Portfolio Analysis & Management (IJPAM).* In December 1995, Professor Schwartz was named the first chairman of Nasdaq's Economic Advisory Board, and he served on the EAB until Spring 1999. He is developer, with Bruce Weber and Greg Sipress, of the trading and market structure simulation, TraderEx (http://www.etraderex.com/). In 2009, Schwartz was named the first recipient of the World Federation of Exchanges' annual Award for Excellence.

Contributors

Karla Amend Clearstream International, Eschborn, Germany

Hugo Bänziger Lombard Odier & Cie, Geneva, Switzerland

Miroslav Budimir Deutsche Börse AG, Eschborn, Germany

Oliver Bussmann Bussmann Advisory, Zug, Switzerland

Andreas Dombret Deutsche Bundesbank, Frankfurt am Main, Germany

Wolfgang Eholzer Eurex Frankfurt AG, Eschborn, Germany

Philipe Enness Global Lead Market Infrastructure, IBM Banking and Financial Markets, IBM Singapore Private Ltd, Singapore, Singapore

Reto Francioni Wirtschaftswissenschaftliche Fakultät, University of Basel, Basel, Switzerland

James H. Freis Jr. Deutsche Börse AG, Eschborn, Germany

Cord Gebhardt Deutsche Börse Group, Eschborn, Germany

Phupinder Gill CME Group Inc., Chicago, IL, USA

Peter Gomber Department of Information Systems, University of Frankfurt, Frankfurt am Main, Germany

Andrew Graham IBM Distinguished Engineer, IBM Banking and Financial Markets, IBM Singapore Private Ltd, Singapore, Singapore

Andreas Grünbichler Wüstenrot Gruppe, Salzburg, Austria

Alexandra Hachmeister Deutsche Börse AG, Eschborn, Germany

Sonali Hazarika Zicklin School of Business, Baruch College, CUNY, New York, NY, USA

Michael Heise Allianz SE, Munich, Germany

Matthias Kluber Deutsche Börse Group, Eschborn, Germany

Alexander Kohler Wüstenrot Gruppe, Salzburg, Austria

Thomas Laux Eurex Clearing AG, Eschborn, Germany

Matthias Papenfuß Clearstream International, Eschborn, Germany

Martin Reck Deutsche Börse Group/University of Basel, Basel, Switzerland

Marc Robert-Nicoud Clearstream Holding AG, Eschborn, Germany

Randolf Roth Eurex Frankfurt AG, Eschborn, Germany

Ulf Schneider Chief Executive Officer, Nestlé S.A., Vevey, Switzerland

Robert A. Schwartz Zicklin School of Business, Baruch College, CUNY, New York, NY, USA

Matthias Stötzel Deutsche Börse AG, Eschborn, Germany

Jan Strecker Deutsche Börse Group, Eschborn, Germany

Stephanie Welkoborsky Genossenschaftsverband e.V., Frankfurt, Germany

Asoka Wöhrmann Deutsche Asset & Wealth Management Investment GmbH, Frankfurt am Main, Germany

Rico von Wyss Bank Vontobel AG, Zürich, Switzerland

Heinz Zimmermann Department of Finance, University of Basel, Basel, Switzerland

Christian Zschocke Morgan, Lewis and Bockius LLP, Frankfurt am Main, Germany

Part I
The Value Chain of Exchange Organizations: Systematic Overview

Chapter 1
Exchange Organizations: Thoughts and Reflections

Reto Francioni

The stock market's animal spirits are symbolized by the striking bronze sculpture of an oversized bull installed in 1989, deep in the heart of New York's Financial District in downtown Manhattan. It is a charging bull, the work of artist Arturo Di Modica. The bull speaks of pure force, unfettered optimism, and aggressive dynamism. And it is quintessentially American, as well as being a grand celebration of Wall Street.

On Frankfurt's Exchange Square, known as Börsenplatz, right in the city center, there is another big sculpture of a bull. It was installed a year earlier, in 1988, opposite the modernized trading floor and venerable headquarters of Europe's most important regulated market. But this bull, designed by Reinhard Dachlauer, is not raging as fiercely as its American cousin. Actually, it has a rather dignified look that might be considered a little static by American standards. But there is an elemental force of nature in this wonderful creature. More importantly, next to this bull, there is Dachlauer's bear, symbolizing the more cautious and more risk-averse phases in trading.

I believe that these sculptures that are on opposite sides of the Atlantic epitomize significant differences between two distinct capital market philosophies and two versions of a liberal social and economic order—the US model being more free-wheeling than the European model. In Continental Europe, this can be seen in those cultures that have matured over hundreds of years around the great trading routes along the river Rhine. A strong vibrant economy has developed here, based on engineering skills, an ethos of professionalism, and an intensive exchange of ideas.

In this light, investments in the real economy are long-term in nature. And they tend to flourish in a culture which rewards careful planning rather than taking an haphazard approach to economic development. In such an environment, capital markets tend to

R. Francioni (✉)
Wirtschaftswissenschaftliche Fakultät, University of Basel, Peter Merian-Weg 6, Basel, 4002 Switzerland
e-mail: reto.francioni@unibas.ch

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_1

play a service role vis-à-vis the real economy. In contrast, Anglo-Saxon capitalism—the model in America—has traditionally been more aggressive in its approach to the capital markets and the real economy. I mean this in a positive sense: it is more open to competition, more skeptical about regulation, and more willing to take risks.

Still, I believe that these two varieties of capitalism are not entirely contradictory. They will perhaps always coexist: euphoria is undoubtedly part of all capital market-driven developments. But risk must be kept in perspective. If forgotten, sooner or later it will come back to cause mischief—in the form of a crisis. Therefore, positive sentiment on markets, the famed "animal spirit" that economist John Maynard Keynes among others thought so highly of, the spirit of Joseph Schumpeter's fabulous entrepreneurs whom the Chicago economist of Austrian birth termed "creative destroyers," is an essential driving force in any liberal society. But it needs to be held in check by a healthy and managed dose of risk consciousness. Different economic cultures find different ways of balancing these two aspects of markets, one emphasizing entrepreneurship, and the other emphasizing risk management. But they never completely ignore each other, at least not over prolonged periods.

Here's another way of looking at the balance between risk management and entrepreneurship—analyzing the relationship between risk and return. The first question: What is the nature of the risk? Secondly, how does risk change in relation to the time period being looked at? And thirdly, investors need to ask themselves if the expected return of their engagement is still sufficient when weighted by the risk inherent in it. No monolithic textbook answer can be devoted to this question, because it depends on individual risk propensities. However, as a general rule, markets work efficiently only if the principle of equivalence holds that the risk inherent in certain assets cannot be passed on from one investor to another without the first being held accountable. Risk and return are, therefore, tied together.

Striking the right balance between risk and return—and, by extension, between risk management and the entrepreneurial spirit—is exactly what exchange organizations do. Therefore, exchange organizations have a place in both types of market economies. Their rules and regulations, corporate governance, and range of products and services are not arcane topics for financial engineers, or "quants." They reflect fundamental decisions about the way we conduct our economic lives and, by implication, how we organize our societies. Exchange organizations support those societies in utilizing the raw energy of market forces in an orderly way—by channeling capital with its pure, unfulfilled potential—into the concrete forms of real investments. How the exchanges are accomplishing this process is the topic of this book.

So what is the actual output of exchange organizations? What do they produce that can be favorably compared with machines built by engineering firms, cars constructed by the automobile industry, or software programmed by IT companies?

Firstly, exchange organizations produce information—information in the form of prices, information about levels of risk and opportunity, about scarcity and abundance in the real economy. This is commonly referred to as an exchange's *price discovery function*. Price discovery is essential for facilitating free trade in an orderly, well-regulated environment. Exchange organizations are like lighthouses for the real economy, offering direction, and doing so without force. Exchange organizations thus put into practice Friedrich August von Hayek's classic argument for market-driven resource allocation over central planning.

An important function implied by price discovery is the valuation of listed companies. The price of a listed company's shares, multiplied by the number of shares outstanding, equals the market capitalization of the company. That's an important measure of a company's economic present value, including its future prospects, as well as its cost of capital. As a consequence, exchange organizations provide a permanent interface between financial markets and the real economy.

Secondly, exchange organizations build, maintain, and grow liquidity pools. Liquidity means, roughly speaking, the immediate availability of assets to be traded against each other, or to be used to facilitate trading. Therefore, liquidity pools are key for all parts of the value chain: in trading, liquidity is essential for efficient price discovery. But similarly so in clearing and custody, liquidity translates into the efficient use of collateral, and ultimately, better risk management at the lowest possible cost.

Thirdly, by formulating rules for equal treatment, especially market access, and by providing transparency for market participants, exchange organizations make sure that risk takers making investment decisions are also responsible for their consequences. These consequences may be a loss attributable to risk, or a reward for having spotted a fundamental opportunity before others. Apart from rules and regulations, transparency is essential for achieving this aim. Transparency refers not only to market prices, but also to the whole central limit order book, which contains buy and sell orders not yet executed. It also refers to essential information on market spreads, depth, and breadth—three important indicators for liquidity.

To that end, exchange organizations ensure that the real economy has access to the financial resources it needs. These resources enable an economy to invest, so it can develop new products and services and, in turn, create employment. For me, managing an exchange organization has never been an end in itself. It has always filled me with pride being able to help make the real economy work better, and to find new ways of achieving this by building gateways to the capital markets.

Risk (and its management) is an essential part of today's diversified exchange organizations.[1] It is because of their commitment not only to efficient markets and investor protection, but also to an equivalence of risk and responsibility, that exchange organizations worldwide operate regulated markets for an increasing range of asset classes. Moreover, they operate clearing houses, settlement and liquidity management engines, market data providers, as well as (last but not at all least) IT companies.[2]

---

[1] I emphasized this already before the onset of the financial crisis; see [2]: "As positive as the general development may be from an investor's point of view, it also contains risks in terms of systemic stability. […] One of the central functions of an exchange organisation—apart from capital allocation, liquidity creation and company valuation—is risk transformation. This refers to a broad range of risks […]: 1. market risks, 2. counterparty risks, 3. operational and transfer risks" (p. 20; own translation from German original).

[2] See, e.g., the interview by *Fortune* magazine with Robert Greifeld, CEO of Nasdaq OMX, 3 September 2014 (Nasdaq CEO: We have to confront brutal reality). Its introduction sums up Nasdaq's development towards horizontal and vertical diversification under his leadership: "When

The evolution of exchanges into integrated financial service providers can perhaps be summed up by three claims, each representative of three different stages of development of exchange organizations in general: "Just another company," "Not just another company," and "Another company."

1.1 Just Another Company

"Just another company" is what a former European exchange leader called exchange organizations right in the middle of the "irrational exuberance"[3] of the 1990s and early 2000s.[4] The phrase, "Just another company," certainly had a rational meaning: The 1990s saw the completion of the first electronic revolution[5] in the world of exchange trading.[6] Electronic order-driven markets, allowing remote access and automatic price discovery at unprecedented levels of transparency, rapidly replaced the old system of intermediation on trading floors. There was the famous "battle for the Bund," the interest rate derivative based on German Government bonds. This battle took place between an exchange of the old type, London-based Liffe, and Deutsche Terminbörse (DTB), the predecessor of what is now Eurex as the proponent of modern, electronic trading. Eurex won because of its technological edge. Apart from bringing about a massive increase in market efficiency, it enabled Eurex to directly connect American traders. Between 1997 and 1999, Eurex's market share increased from 30 % to nearly 100 %.[7] Eurex is, of course, only one example of many. The pioneer was Nasdaq, with its electronic quotation system dating back to 1971.[8] And in 1977, the Toronto Stock Exchange pioneered electronic trading among national stock exchanges.[9] Electronic trading also had enormous consequences for the governance of exchanges—a case of superstructures changing as a result of developments at the economic base: The old systems—exchange participation limited to a presence on the trading floor—became obsolete. Trading and exchange ownership could now be separated. By the early 2000s, this paved the way for exchange organizations becoming publicly listed entities, and, in turn, entering

---

Robert Greifeld, 57, became Nasdaq's chief in 2003, it ran one equity market in the U.S. Today, it owns and operates 26 markets globally for trading stocks, bonds, derivatives, and commodities; its technology runs 70 markets on six continents."

[3] See [8].

[4] For the narrative justifying the thinking behind this, see [4].

[5] Strategically, IT along the whole value chain means faster, cheaper, and better performing systems; the game of economies of scale, with a base of fixed cost; international/global reach; more transparent trading, and better surveillance; increasing consolidation pressure; and that liquidity pools can be built faster and at lower cost, but they are much more vulnerable.

[6] See [6, 7].

[7] See [3].

[8] See ([6], 52f).

[9] See ([6], 84ff).

a new dimension in leverage for global geographic expansion, and vertical and horizontal integration and consolidation. First, that occurred within one economy, then within one time zone, and later on a global scale. In a way, the bulk of exchange organizations had, in fact, become "just another company."

1.2 Not Just Another Company

Starting around 2005, another paradigm shift took place in exchange strategies. This change partly anticipated, and was then strongly accelerated, by the global political response to the financial crisis of 2007ff. This response is nothing less than a U-turn in the policy approach to the financial sector: from deregulation to reregulation. When embedded in diversified exchange organizations, regulated markets are a significant part of the solution by policy makers and market practitioners worldwide. This is because, at its core, central counterparty clearing has been identified as having the potential to substantially change and control the risk structure of the entire value chain, including the OTC markets. And, thereby, central counterparty clearing can strengthen the systemic stability of capital markets. This change and concentration of risk management for each and every market by a clearing house leads me to the central thesis of this chapter, and of the entire book: in the new order of capital markets, the diversified exchange organizations will be able to provide solutions to the new challenges, as the far-reaching response for overcoming the endemic instability of the old deregulated world. What is more: these permanent risk management solutions are useful for regulators, banks, and the real economy. In this sense, each exchange organization is "not another company." By being a neutral arbiter of conflicting interests, it is different from other companies, banks included. And each exchange organization is differentiated by creating and organizing capital market infrastructure where safe and orderly trading is possible, and the equivalence of risk and responsibility is reestablished.

It seems quite clear that the increasingly deregulated and highly leveraged markets in the years before the crisis of 2007ff did not perform as efficiently as some economic theorists had once believed. But does this mean that we should abandon, as some proponents of behavioral finance seem to suggest, the idea of markets as instruments of rational decision making altogether? Definitely not!

What we need is a form of regulation that encourages rational decision making, keeping a close watch on risk and return on the one hand, and investor and system protection on the other hand. Freedom and regulation do not need to be at odds with each other. Regulation that is both efficient and effective provides the framework for competition that is free insofar as no participant enjoys an unfair advantage over any other participant. This may indeed sound idealistic. And, in reality, we probably need to confine ourselves to approaching this ideal in the best possible way, without ever reaching it. We will continue to strive for a permanent optimization of regulation, in quality and in quantity. This means that we must stop seeing regulation and free markets as a contradiction. Not to put too fine a point on it: Only regulated markets are free markets—with the important proviso, however, that regulation needs to refrain from intervention in the free interplay of supply and demand. Regulation must neither interfere with price discovery nor predetermine market outcomes. But it needs to define the rules for fair price discovery and capital allocation. Unregulated markets, on the other hand, are free only for those who have privileged access to either information, capital, or manpower—or to all three combined.

This argument has one important implication: There must never be a market for regulated markets themselves. The legal system of a liberal society rests on cultural or even—as Immanuel Kant would argue—on a priori rational principles it cannot guarantee by itself. Similarly, the principles underlying the economic system of a liberal society—rest on legal and cultural foundations that do not just emerge spontaneously from the interaction of self-interested individuals in a Hobbesian state of war.[10] These principles include equality of market access and information (especially full availability of price-sensitive information), absence of market manipulation, freedom from distortion by the abuse of quasi-monopoly power, or undue government intervention. These same principles need to be held alive in public debate. They need to be secured by a market-friendly legal system, and implemented by institutions committed to organizing markets following these principles by their legal setup and business model, in other words by exchange organizations.

This does not mean that these exchange organizations should be completely exempted from competition. A bout of competition between globally operating exchange organizations, as well as supervised competition with other less regulated platforms, is useful in fuelling the "animal spirits" and sensitivity to customer needs. But it must be noted that opening up "*dark pools*" in the well-lit world of the market will be self-destructive, as the crisis of 2007ff should have abundantly demonstrated.

Simply put, the new order calls for a rediscovery of the virtues of adequate market regulation. Mind you, it should support a free market—regulation does not mean bureaucratic excess. The problem is the wrong regulation, or overregulation, not in regulation itself. Let me explain. Regulation means fairness, transparency, and equal opportunities for all, competitors included. This rediscovery should be welcomed—at the same time as maintaining a strict limit on its potential costs. The rediscovered virtues underlying adequate regulation include:

---

[10] The Penguin Dictionary of Philosophy ([5], 2nd edition, pp. 278–279): "In pursuing felicity as they see it, people naturally exercise the right—'the right of nature', Hobbes call it—of judging for themselves how best to get what they want. Problems arise when individuals want the same thing, or when greedy or vainglorious individuals—they need only be a minority—act in character and want more goods or esteem than their neighbours. In all of these cases, commonplace in the nature of things, people are anti-social. They come into conflict. The conflict need not manifest itself in outright fighting, but there is always a danger that it will. Indeed, the right of nature entitles people to use violence in pursuit of their aims if they judge it to be appropriate. Even extreme violence may be justified by the right of nature. In this way the state of nature can amount to a state of war, and indeed is likely to. Either violence will be resorted to gratuitously by the greedy and vainglorious; or it will be resorted to reluctantly and reasonably by moderates intent on protecting their lives and goods from those who are immoderate."

- Safety and individual responsibility for risk taking
- Integrity and the avoidance of excessive exposures
- Efficiency and transparency, brought about by a simplified market structure
- Last but not least, a new emphasis on burden sharing: participation of the financial industry in the costs imposed by a crisis on government and, ultimately, tax payers

Two overriding principles, however, remain indissoluble: investor protection and system protection. In other words, fair and equal treatment of each market participant, as well as rules, regulations, and technologies that guarantee systemic stability, is an enduring principle.[11] It is not a coincidence that they are also at the core of exchange organizations' objective.

The new emphasis on the values of safety, responsibility, integrity, efficiency, transparency, and burden sharing has led to a number of new regulatory initiatives. What is remarkable is that they harken back to a truly global initiative: The G20 Pittsburgh Summit of 2009 (the gathering of leading economic powers worldwide) argued for clear standards for transparency and risk management of the world's financial markets. Global imbalances are perhaps inevitable in implementing these measures in a world still divided by economic, political, and cultural divergences.

1.3 Another Company

The new global Basel III capital requirements from 2013, the passing of the European Market Infrastructure Regulation EMIR in 2012, and the US Dodd–Frank Wall Street Reform and Consumer Protection Act in 2010 opened a new era of regulatory change aimed at re-embedding the financial sector more firmly into society. While it took further time to pass the implementation measures to apply the new regulations to market reality,[12] new needs were also created for exchange organizations to find answers for market participants in the areas of liquidity, capital efficiency, and collateral management.[13]

Apart from the promise of greater systemic stability and better investor protection, the new regulations led to new burdens for market participants. That was especially so in the higher capital requirements to deleverage their balance sheets. As a consequence, new customer needs have materialized from an exchange organization's perspective: liquidity has become an even more pressing issue than before. More capital efficiency has become important in the drive to squeeze more liquidity

[11] Referring to both IT and market stability.

[12] The sheer scale of the Dodd-Frank Act is as impressive as it is severe in its impact on the market: The 900 pages of the Act itself are complemented by further 9000 pages of implementation measures. By way of comparison, the user manual for the space shuttle only has 1200 pages.

[13] For example, the Dodd-Frank Act (http://www.gpo.gov/fdsys/pkg/PLAW-111publ203), Basel III (www.bis.org/bcbs/basel3.htm), or the EU Capital Requirements Directive IV (ec.europa.eu/finance/bank/regcapital/legislation-in-force).

out of limited asset pools. This enhanced efficiency also makes it easier to fulfill new requirements in the realm of collateralization. An exchange organization, understood as "another company," needs to offer answers to these new needs in the emerging new capital market order.

Exchange organizations must accomplish three things to adjust and to assist capital market participants in responding to this new order: First, they must diversify their business models, in the process appreciably lowering their exposure to now mature products and services with intensive competition, while expanding into higher growth areas, such as innovative derivatives based, e.g., on dividend payments or volatility.

Second, exchange organizations should add elements that give them (and both their users and their owners) more stability during periods of adverse market conditions. For instance, this could be provided by central counterparty clearing and liquidity management services. Third, exchange organizations should exploit their in-house know-how in relation to IT and, most important, assess the impact of digitalization[14] on the entire value chain strategy for the business model of the future.

The future business model of exchange organizations and even the whole financial industry will be heavily affected by the evolving fintech (financial technology) industry.[15] The fast growing fintech start-up scene is mainly driven by regulatory encouragement and changing customer needs, but increasingly also by competition. It will affect the whole value chain of securities business—in part and as a whole—which means investment management, trading, information, and risk management and administration. In the trading area it will digitalize stock, bond, and derivative trading features. And with respect to information, access, and handling, fintech will be key in connecting the owner of the order flows—buyers and sellers—indirectly or directly. This impact also includes analytics and management of big data, collateral and portfolio management, ex post trading decisions, performance steering and measures, FX retail and institutional trading, and payment systems. Fintech will also underpin and challenge the strategic role of data with respect to customers and markets. That is why fintech has a high potential of innovation AND disruptiveness: the extent and the magnitude are open. Fintech will challenge the role of any intermediary function within not only the value chain of exchange organizations but also their customers. And it will therefore shape business models, long-term structure of the industry, regulation, and customer behaviors.

---

[14] Digitalization refers not only to the use of IT systems to organize markets, but has even wider implications for the economy as a whole. It has the potential to destroy and redefine value chains from scratch. Under labels such as "Industry 4.0" in Germany, digitalization will further enhance the role of IT by automating decisions about production, marketing, and distribution to an extent unknown before. In exchange trading, the use of algorithms for automating trading decisions has to some extent anticipated this development. Of course, this also means that IT-driven processes might become more vulnerable to manipulation or to organized cyber criminality. In response, risk management processes need to be supplemented by IT security plans.

[15] For details see WEF Report on the future of financial markets: http://reports.weforum.org/future-of-financial-services-2015/.

Regarding IT in general, the key players of today should evaluate in detail if, where, and how they have to adapt and implement the new paradigms: For exchange organizations "blockchain" could play a crucial role, because blockchain technology [16] is based on a distributed peer-to-peer network. This allows for a certification of ownership and clearing of transactions without a central repository or the interaction of a central administrator: users can interact directly without any intermediary. Each transaction between members of the network is verified and validated directly to guarantee a valid transaction between two individual accounts and to avoid the risk of double spending or counting.

Diversified exchange organizations are particularly well qualified to support both regulators and customers in implementing and adapting to the "new order."

Diversification has two dimensions: On a horizontal level, it means, first, adding new asset classes to existing offers, and, second, expanding geographically. On a vertical level, it means, first, building integrated value chains; second, making the elements of these chains interact; third, developing new and innovative products and services from this interaction; and fourth (especially in the light of digitalization), adding customized and tailor-made products and services.

Therefore, exchange organizations need to morph into comprehensive financial infrastructure providers, genuinely becoming "other" companies. The most important of these are liquidity hubs offered by post-trade service providers, and clearing services for OTC derivatives. These two are closely interlinked: OTC clearing is a response to the need for collateralization that improves the systemic stability of markets. It is liquidity management that provides answers to the need for an efficient use of this collateral to make the pressure on scarce capital resources manageable both for banks and, even more importantly, for the real economy served by the financial sector.

1.4 Making Risk Management Work: OTC Derivative Clearing

According to the Bank for International Settlements, by the end of 2013, the volume of OTC derivatives markets, measured in notional amounts outstanding, amounted to more than US$700 trillion.[17] So far, the percentage traded on derivatives

[16] Blockchain is a database with the following attributes:

- Modular and distributed with a network of communicating nodes, providing a distributed and public ledger accessible to whom it may concern.
- Transactions are stored and verified on every node of the network and each copy contains the full history of all transactions.
- Consisting of blocks that are added in chronological order, building the chain. Each block contains a certain number of transactions and is inscribed with an automatically generated check number.

[17] http://www.bis.org/statistics/derstats.htm.

exchanges and cleared via CCPs is very small in overall trading. This means that, up to the present day, a huge amount of extremely complex financial instruments that support highly leveraged trading strategies is totally unregulated and unsupervised. The time has come to change this, segment by segment. And the first steps towards that change have already been taken, at least in the USA, with the implementation of the Dodd-Frank Act. Europe, in contrast, is still catching up, with an ominously sounding regulation called EMIR, short for European Market Infrastructure Regulation.

The burdens on banks because of new capital rules are significant: According to estimates by the Basel Committee on Banking Supervision as of April 2012, banks worldwide are facing an aggregate shortfall of stable funding of €2.8 trillion—approximately US$3.7 trillion—in fulfilling the additional liquidity requirements of Basel III. In addition, the new clearing obligations, although improving systemic stability, pose new challenges to the liquidity management of banks. Therefore, there is an urgent need for financial infrastructure providers to assist banks in managing these resources efficiently.

CCP clearing deals with the three chief causes of the financial crisis: the decoupling of risk assumption and responsibility; the insufficient collateralization, especially in OTC derivatives trading; and the complexity caused by the interconnectedness of market participants.[18] CCPs do not take risk themselves; they manage risk on behalf of their members as they:

- Calculate the risk of each open position, ideally in real time, and thus make risk transparent.
- Produce the equivalence of risk and responsibility by collecting collateral in the form of margins.
- Build additional lines of defense by charging initial margins.
- Have a strong incentive to manage risk carefully because they also inject substantial amounts into the reserves accumulated (in case of an emergency) from their own resources.

These are the measures they take to tackle the first two causes of instability which led to the 2007 crisis. They also substantially mitigate the potential domino effects resulting from interconnectedness: clearing also includes "novation"—the legal stepping between each trade. This enables multilateral netting and vastly reduces the number of transactions which need to be cleared.

All in all, clearing makes markets more efficient and more stable. In the final analysis, this creates massive net benefits for the broad economy, although it also increases the pressure on scarce collateral. In this way, it adds to the demands on banks by regulatory reforms that insist on increased capital requirements—the other important way of better managing systemic risk in the future.

[18] For this and the following, cf. [1]. Deutsche Börse's successful General Collateral (GC) Pooling service for the interbank market, which also includes the Eurex central counterparty, sets an example that CCP services can be extended beyond securities and derivatives trading; see http://www.eurexrepo.com/repo-en/markets/gc-pooling-market.

1.5 Making Risk Transparent and Affordable: Global Liquidity Management

Banks and insurance need to deal with the greater capital requirements introduced by Basel III and *Solvency II*, respectively. That's on top of the additional requirements on collateralization in derivatives trading, a combination that creates massive pressure on the liquidity of the financial sector at large. And, by implication, it leads to a massive shortfall for the real economy. Collateral cannot be created out of thin air. It is a scarce resource. Therefore, the only solution for managing this shortfall is the more efficient use of this resource.

Once again, well-diversified exchange organizations have in recent years developed new products and services that, in facing this new challenge, help not only banks and brokers, but also buy-side firms such as insurance companies and even the real economy. The solution is access to a broad diversity of assets, processing management across the full value chain, and doing so with a global reach available to customers.

The post-trade business, traditionally operated by organizations working independently of securities exchanges, diligently processing myriads of transactions, has gained substantial importance in recent years. Being less affected by the business cycle than trading and clearing, it can provide stable earnings to exchange organizations. Also, because of the operational safety it provides and the high level of confidence it creates, it helps build stable bridges and unbreakable vaults in the labyrinthine networks of trading on regulated markets and other venues. But the main reason for the ascent of the post-trade business is the increasing demand for collateral, which can be turned into liquidity as the need arises. Collateral and liquidity management has became, next to trading and clearing solutions, the third pillar for exchange organizations seeking to provide stability and transparency to the financial sector.

Exchange organizations are champions of intelligent markets. However, such markets will stand the test of time only if the freedom they offer is wisely regulated. In today's world, the solutions for such wisely regulated freedom are increasingly shifting from trading (the focus of the last century) to clearing and post-trading facilities. This is why regulated markets with a diversified value chain are best able to fulfill the demands of today. They can do this for their owners, customers, the real economy, and society at large. The order of the day must therefore be to unleash technological innovation and entrepreneurial zest in strengthening and extending such markets—in a spirit of fairness and trust.

In farming, bulls (not of the iconic kind in New York and Frankfurt!) are required to perpetuate fertile livestock. However, a farmer who would knowingly let a bull rage at random would be considered foolish—and rightly so. His livestock business would soon collapse. Bulls take risks all too easily, and the idea of personal responsibility is alien to them. It is the farmers caring for them who need to act responsibly. In today's world of electronic business, we are all still in the debt of the early farmers who formed the first societies, having to face the most elementary risks that threatened their very survival. Being challenged today in this way, and enduring the most elementary risks by acting together with dignity, is a superb school for one's own character. It teaches an excellent lesson—to act responsibly.

References

1. Eurex Clearing / Deutsche Börse Group (2014): How central counterparties strengthen the safety and integrity of financial markets. Frankfurt / Main: Deutsche Börse AG; 35 p.
2. Francioni, Reto (2007): Börsen als Fundament funktionierender Kapitalmärkte. In: Strebel-Aerni, Brigitte (ed.): Internationale Finanzmärkte zwischen Effizienz und Sicherheit. Kapitalmärkte: Herausforderungen & Chancen. Zürich: Schulthess; pp. 19–25
3. Gerstenschläger, Frank (2002): Elektronische Börsenstrukturen im globalen Finanzmarkt. In: Krumnow, Jürgen / Gramlich, Ludwig / Lange, Thomas A. / Dewner, Thomas M. (eds.): Gabler Bank Lexikon. Wiesbaden: Gabler; pp. 236–241
4. Mattern, Frank / Seifert, Werner G. / Streit, Clara C. / Voth, Hans-Joachim (1997). Aktie, Arbeit, Aufschwung. Wie der Finanzplatz Wirtschaft und Gesellschaft wieder in Schwung bringt. Frankfurt / New York: Campus; 203 p.
5. Mautner, Thomas (2005): The Penguin Dictionary of Philosophy. Suffolk: Penguin Books.
6. Schwartz, Robert A. (1991): Reshaping the Equity Markets. A Guide for the 1990s. New York: Harper Business; 452 p.
7. Schwartz, Robert A. / Francioni, Reto. (2004). Equity Markets in Action. The Fundamentals of Liquidity, Market Structure & Trading. Hoboken: Wiley; 468 p.
8. Shiller, Robert J. (^{2}2005; ^{1}2000). Irrational Exuberance. Princeton et al.: Princeton University Press; 304 p.

Chapter 2
An Exchange and Its Value Chain

Reto Francioni

2.1 Introduction

This book covers **price discovery**[1] at **stock exchanges** with reference to both the **primary** and the **secondary market**. Markets are constructed as liquidity pools and, therefore, the main task of a stock exchange is twofold:

1. Build, maintain, and grow liquidity pools in each **listed stock**: *The fight for **liquidity**!*
2. Enable fair and adequate price discovery in every listed stock during the entire trading time (Fig. 2.1).

   This is facilitated by:

- Customer-oriented market structures
- Fair rules and regulations
- Market-oriented operations of the exchange
- Highly competitive technology
- Effective surveillance
- Efficient fit into the value chain, which covers **clearing**, **settlement**, and **custody**; the customer interface; exchange members; **brokers** and **dealers**

Price discovery is a key function of the market structure, and it requires a tailor-made, customized solution for each product class to attract as much liquidity as possible (Fig. 2.2).

The value chain of a stock exchange has many elements:

[1] Cf. Glossary, which is an integral part of this book: every term and every abbreviation in **bold** are explained in the **glossary**.

R. Francioni (✉)
Wirtschaftswissenschaftliche Fakultät, University of Basel, Peter Merian-Weg 6,
Basel, 4002 Switzerland
e-mail: reto.francioni@unibas.ch

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_2

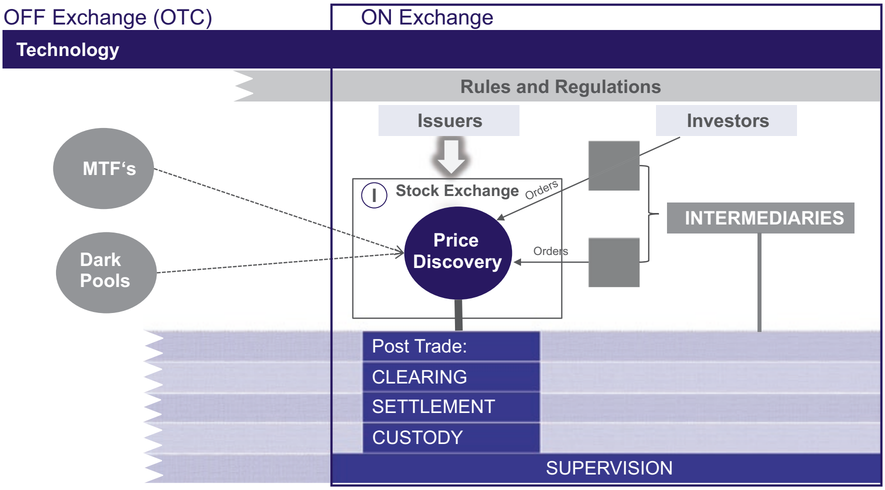


Fig. 2.1 The exchange landscape: overview

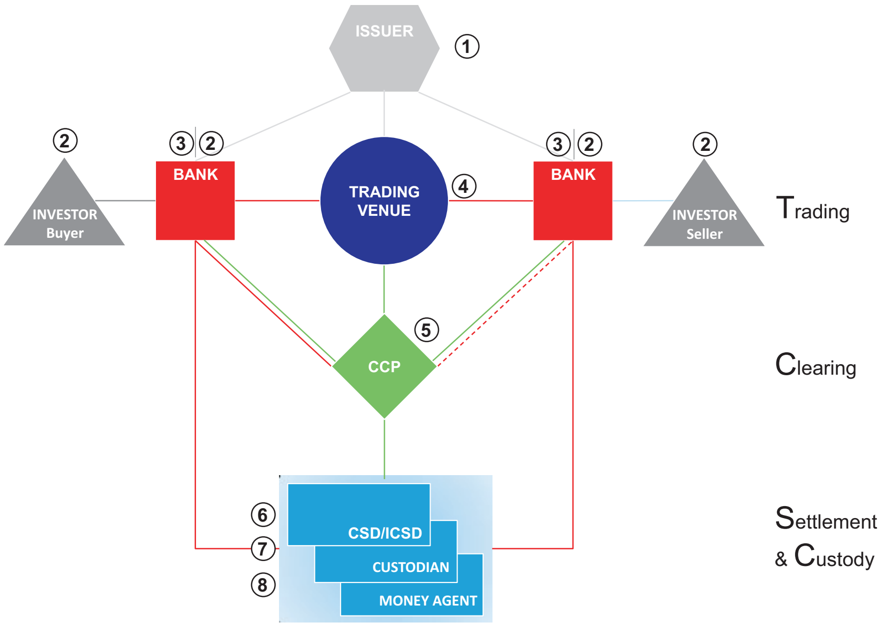


Fig. 2.2 Elements of the value chain

1. Product creators are called issuers. An issuer sells all the shares or at least a certain percentage of the share capital (depending on the listing rules of a stock exchange) to private investors through a public offer, a so-called **initial public**

offering (IPO). By placing stocks in the primary market at/through a specific stock exchange, they are listed and therefore are ready to be traded at the stock exchange in the so-called secondary market.
2. The flow and order creators: These include intermediaries, if they trade from their own account (sell-side or principal dealer); institutional investors (called the buy side) like hedge funds, insurance companies, or retail investors; and private persons buying and selling stocks.
3. The order routers: These are all stock exchange members, called intermediaries, who are acting as brokers.
4. The trading venue or the stock exchange where the central order book is monitored and the price discovery takes place.
5. The clearing house/central counterparty (CCP): The CCP mitigates counterparty and market risk, reducing costs through netting.
6. The settlement organization/central securities depository (CSD): The CSD ensures the delivery and payment of an executed order that is hedged in time, quantity (size), and quality.
7. The custody organization: Covering the administration and safekeeping of stocks.
8. The money transfer system: Facilitating the post-clearing money flow.

A buy or sell order starts with an investor before it has to be handled by an intermediary and routed to the exchange where the price discovery and therefore the actual trade take place.[2] The trade is then routed to the clearing house, where counterparty and market risks are mitigated, and the **netting** of trades takes place. Eventually, the execution of a trade is finalized and confirmed through the settlement organization. The settlement (**delivery vs. payment**: **DvP**) includes custody processes that ensure the delivery of a security, and the payment system, covering the money flow.

This process, which starts with the initial order and ends with the confirmation of the trade to the investor who bought or sold the respective stock, is called the *value chain*.

Traditionally, the trading side includes the clearing part, which is why settlement is often called the post-trade part of the value chain. But since trading and clearing are different organizations (in legal, regulatory, and surveillance terms) and have different functions, they are treated separately from each other. Therefore, everything following clearing in the value chain is defined as *post-clearing*, while at the same time everything following trading in the value chain is defined as the term *post-trading*.

[2] A trade takes place when a buy order and a sell order of the same stock are matched and executed, identically in price, time, size, and place.

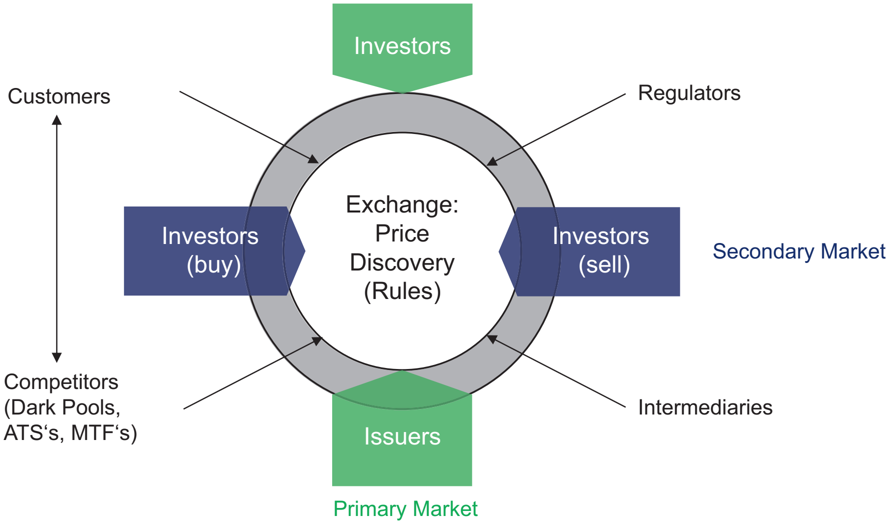


Fig. 2.3 What is an exchange?

2.2 The Exchange: Link to the Real Economy Through Investments

2.2.1 Primary and Secondary Market

A stock exchange is *the* primary liquidity pool in a stock, when its price discovery takes place (Fig. 2.3). Trading and price discovery on a stock exchange is called a secondary market since securities/stocks are already issued. Therefore, price discovery takes place with a given quantity of shares (namely **free float**), and it leads to changing prices per share based on the relative selling and buying pressures of investors. It is a neutral, regulated, and fully transparent marketplace for stocks:

- Neutral, because there is no conflict of interest, and equal treatment of the members and issuers is secured.
- Regulated, because there are laws and other regulatory acts that define how to run a stock exchange, how to handle listing and trading, and how to protect investors and the system as a whole. To enforce rules and regulations, a market supervisory authority is tracking the price discovery at the exchange itself (this is called market surveillance) as well as outside of the exchange on a national level. For instance, in the USA this is the Securities and Exchanges Commission (SEC), in the UK it is the Financial Conduct Authority (FCA),[3] in Singapore it is

[3] The Financial Conduct Authority (FCA) was formed in 2013 as one of the succeeding agencies of the Financial Services Authority (FSA) in the UK. It is a quasi-governmental agency that regulates

the Monetary Authority of Singapore (MAS), in Germany it is the Federal Financial Supervisory Authority (BaFin), and in France it is the Autorité des marchés financiers (AMF).

- Transparent, because there is an open, centralized **order book** for market orders (**COB**) and/or limit orders (**CLOB**). Market participants and—time-delayed—investors are offered an insight view of, and information about, orders in the book for every traded stock: *You see what you get and you get what you see*.

Neutrality, regulation, and transparency are essential basics for the market integrity of an exchange. Market integrity will eventually lead to trust in a market and a financial system: The higher the market integrity, the higher the trust. Market integrity is a key asset of an exchange and of the financial value of an economy, because there is, and there has to be, a close interrelationship between a capital market and its corresponding real economy. In essence, a capital market has to serve the real economy and not vice versa. In this respect, a stock exchange and its value chain are the link between capital markets and the real economy. Similarly, for cleared OTC products, the clearing house is another link to the real economy.

Differentiated from the secondary market, where price discovery takes place via the matching of investors' buy and sell orders, is the primary market. By means of an IPO and as a result of a stringently structured listing procedure, the issuer publicly offers all or parts of its shares to investors. Hence, the primary market is a syndication-supported placement of new shares at an exchange to public investors by issuing companies. If a company offers all of its shares, the free float of the listed company is 100 %. Depending on the listing segment, the free float should be at least 20 %. In both the primary and the secondary market, capital allocation takes place, thereby linking capital markets to the real economy. Due to this fact, price discovery is a process for both markets.

In the primary market the issuing price has to be discovered in an auction-like procedure. Then, as a link to (and already part of the secondary market) the first price has to be fixed at the exchange (also by means of an auction-like procedure). Following this, the secondary market is established and price discovery takes place in the CLOB. It does so in order-driven markets, either through periodic auctions or continuous trading (Fig. 2.4).

Trading in the secondary market can be carried out using different platforms:

1. On-exchange trading describes the *traditional* form of trading by means of a regulated exchange organization. Benefits of this trading form include supervised rules and regulations that guarantee system and investor protection, as well as transparent processes of price discovery within a central order book.

retail and wholesale financial service provider. The second agency formed is the Prudential Regulation Authority (PRA). As part of the Bank of England it oversees the regulation and supervision of banks, investment firms, etc.

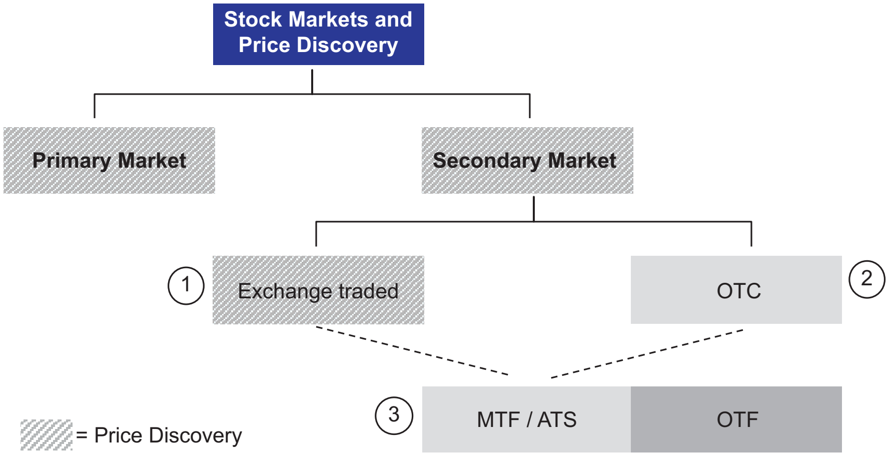


Fig. 2.4 Price discovery on exchanges

2. Off-exchange or OTC trading describes direct and relatively unregulated trading between two dealers by means of electronic trading systems. Since there is no neutral and supervised mediator between buyer and seller, and because order books are usually not accessible, no regulated price discovery takes place and transparency is limited. This form of trading bears considerably more risk than on-exchange trading.

   The OTC market compared with regulated markets may be generally characterized by:

- Fragmented liquidity
- Inaccurate price discovery and negotiation
- Limited transparency and acceptability
- Operational inefficiency
- Expensive and slow transactions

(3) **Multilateral trading facilities** (MTF, EU)/**alternative trading systems** (ATS, USA) are non-exchange trading venues with electronic systems that are operated either by an exchange organization or other market operator. Investment services are provided by bringing together third-party buying and selling interests in the system under nondiscretionary rules resulting in contracts. **Organized trading facilities** (OTF, EU) is a term described by MiFID II[4] that subsumes all trading venues that are neither a regulated market nor an MTF. This category was introduced in order for MiFID regulations to cover *all* investment services that bring together third-party buying

[4] MiFID is short for the European Markets in Financial Instruments Directive, in force since 2007, that is currently under review. The EU Commission's proposal for a revised directive is hence called MiFID II.

and selling interests in financial instruments in a system to form a contract (especially **dark pools**).

2.2.2 Mission: Capital allocation through price discovery

2.2.2.1 Precondition: Build, maintain, and grow liquidity pools

- Macroeconomic objective: Fair capital allocation that links the real economy to capital markets

Every economy has to have a high-performing and reliable capital market structure to facilitate growth and wealth. An exchange is an essential link between the real economy and the capital market, financing growth of the listed companies through a transparent, regulated, and neutral procedure in the primary (IPO) and in the secondary market. Exchanges and their post-trade organizations reduce the **systemic risk** of markets and enhance market integrity, largely in terms of transparency, security, and stability. The neutrality of exchanges is the key to supporting fair trading and equality for all market participants.

- Microstructure objective: Effective and efficient price discovery
- Legal objective: Investor and system protection
- Operational objective: System reliability

Reliability refers to the IT trading platform with respect to availability, latency, and capacity. The service availability for a trading platform is measured by the total online time of a round trip, with all components (trading, clearing, settlement) working flawlessly. All incoming orders are time-stamped at the interface between member and exchange, and all trades are time-stamped at price discovery and as they pass back from the exchange system to the member system. All components are usually monitored 24 h and have a backup component that can take over within minutes. For Deutsche Börse Group IT, for example, the service availability for customers considering all components (e.g., network, back-end services, and application software) is considered to be *ok* if the availability is higher or equal to 99.974 % (approximately 1 h per year with 257 trading days).

Latency defines the time for a roundtrip of a **transaction**. Every transaction routing, storage, or execution of **matched** and unmatched orders is time-stamped along with their complete itinerary for the entire **straight-through processing** (STP). The latency for access to trading is the time measured from the technical interface to the member to the entrance in the COB. **Host** and **local backbone latencies** are measured round trip times between member interface and host. Latency is also measured as the time for the access of a member to the COB (e.g., leaving the interface still placed in the COB and ready to be matched).

Fig. 2.5 Price discovery process in the secondary market

Capacity is the ability of an IT platform to handle a specified load, a given percentage above the highest observed daily peak. The percentage should include foreseeable changes of the system load, as well as business volume estimates due to expected market developments (e.g., changes in rates, or due to new products or services).

2.2.2.2 Price Discovery

After a company (the issuer) is publically listed and admitted through an IPO in the primary market, the stock of this company is traded publically at the exchange in a central order book (COB): This is the secondary market. In the secondary market investors value the opportunity to participate as co-owners in value creation for a company. Secondary market means trading and trading means the execution of an investment decision. An investment decision is eventually made by the owner of the asset and therefore the order flows (Fig. 2.5).

Buyers and sellers meet in the central order book. The central order book may be transparent or closed; if the latter, this means no insight for investors or intermediaries. The transparency of a COB may differ in:

- Time (data postponed for the public)
- Scope (accumulation in all or bid by bid, and ask by ask, anonymous with market participants)
- Segment

All buy and sell orders for a stock are collected, brought into the central order book, and either executed or stored in the book. The matching of the bid and ask sides takes place, thereby creating the execution price and the trade. The overriding principle is always to match as many orders as possible. There are different matching procedures to get this result:

- **Auctions** with predefined algorithms
- **Market making** with predefined rules

When a market is opened, it changes its mode from *opening* into *continuous*. In **continuous trading**, orders are executed against the spread or, if no execution is possible, stored in the COB. A COB contains market orders and limit orders;

if it contains limit orders only it is called central limit order book (CLOB). An IPO or the primary market can also be viewed from a price discovery angle: The evaluation or fixing of the first price starts the secondary market by applying a special *price discovery procedure*.

2.2.3 Performance Criteria

The following performance criteria set the stage for liquidity creation:

- Effectivity[5]: High quality
  - Optimal market structure: Customized and tailor-made
  - Competitive regulation: Ockham's razor[6]
- Efficiency[7]: Low cost
  - Information
  - Clearing
  - Transaction costs: Implicit and explicit, market impact, commissions, margin, fees, investor costs, clearing costs, surveillance costs
- Integrity: Trust—Protection of investors, system protection and protection of functionality, fair price discovery, protection against insiders, market transparency, data consistency
- Reliable surveillance
- Attractive product range: Products and markets, single and innovative segmentation
- Distribution power through the trading platform
- Neutrality, no conflict of interests and equal treatment
- Robust and performing technology (volume and time)

2.2.4 Home Markets

A home market is the domestic stock exchange for domestically listed companies, and it is usually the biggest single liquidity pool in this stock. German companies (i.e., Daimler, Deutsche Bank, Lufthansa) are listed on the main German exchange Frankfurter Wertpapierbörse (FWB) and traded on its principal trading platform XETRA.

Elements of a home market are (generally):

---

[5] Do the right things!

[6] Ockham's razor is a principle of logical problem solving stating that out of a selection of hypotheses, the simplest or the one with the least assumptions should be used to proceed.

[7] Do the things right!

- Biggest single liquidity pool in domestic stocks
- One law per subject
- One set of accounting principles
- One market authority and surveillance
- One tax regime
- Local investment community
- Local banking community
- A single STP that covers the whole value chain
- Link to domestic real economy

2.2.4.1 The Home Market Principle

Every cash market is the *(world) champion* in turnover and therefore in price discovery for its domestically listed stocks (e.g., SIX for Swiss stocks, NXT for French stocks, Frankfurt XETRA for German stocks). This is especially the case for **Blue Chips** (e.g., France: CAC 40, UK: FTSE 100, Germany: DAX 30, Switzerland: SMI).

Because the home market is the biggest liquidity pool in a domestic stock, everybody is trading against it as a reference market. For a stock, this is necessary, because the reference to the home market price is a component of the secondary market strategy of a (listed) company.

2.2.4.2 Reference Functions

Reference function in the EU:

- The post-MiFID situation in trading securities is coined by the fragmentation of market liquidity to exchanges, MTFs, dark pools, and other off-exchange trading venues.
- Home markets are used as a reference source for price discovery on third-party trading venues.
- Off-exchange vendors as well as exchange-traded retail offerings are often focused on limited product offerings.
- Regulatory differences allow for a wide range of trading models even in exchange offerings, while substantive trading surveillance is sometimes questionable.
- Flow providers (retail) are looking for incentives to control the flow of orders to certain places/venues; different models have been established on-exchange as well as off-exchange.
- Institutional flow provides the main share of sales volumes on the trading platform. High liquidity, neutrality, transparency, and execution quality are principal reasons for the leading market position of a trading platform.

Reference functions in microstructure are for example price discovery in the primary and secondary market, price reference for indices/ETFs/mutual funds/

structured products/futures, and options or OTC markets/trades for MTFs, dark pools, and other off-exchange trading venues. Details are described in course of Chap. 4.

2.3 Layers of an Exchange

A layer of an exchange describes the complete scope and contents of an exchange under a specific point of view or criteria. In this book an exchange is described from four different points of view, namely:

1. Organization
2. Technology
3. Regulation
4. Stakeholder

2.3.1 Organization

Every exchange has its corporate governance structure. If listed, this governance has to fulfill listing criteria and conditions. In addition, especially in Europe, there are still exchange councils regulated by public law.

In addition to its corporate structure, every exchange has an organizational module in the following areas: market surveillance, market supervision and control, order and trade management, information dissemination, and master data.

2.3.2 Regulation

1. The principle of legal empowerment: Eventually, every legal action needs an underlying law; for example, in the USA the legislative power lies with congress as derived from the US Constitution (see Article I, Section I). Penal code (Fig. 2.6).
2. The legal principle of competence delegation: The delegation of power to a regulated exchange by means of national or federal legislation (a securities or stock exchange law) allows the exchange to set its own rules and regulations. The law in question must explicitly delegate this power to establish rules. The competencies must be clearly specified and covered by the delegating legislation.
3. The principle of self-regulation: For example, in the USA, securities exchanges are required to register with the Securities and Exchange Commission (SEC) as laid out by the Exchange Act.[8] The Act also lists the requirements exchanges must meet in order to register with the SEC, including the obligation to provide a set of rules that regulate the conduct of their members (for example, listing and

[8] http://www.law.cornell.edu/uscode/text/15/chapter-2B [1].

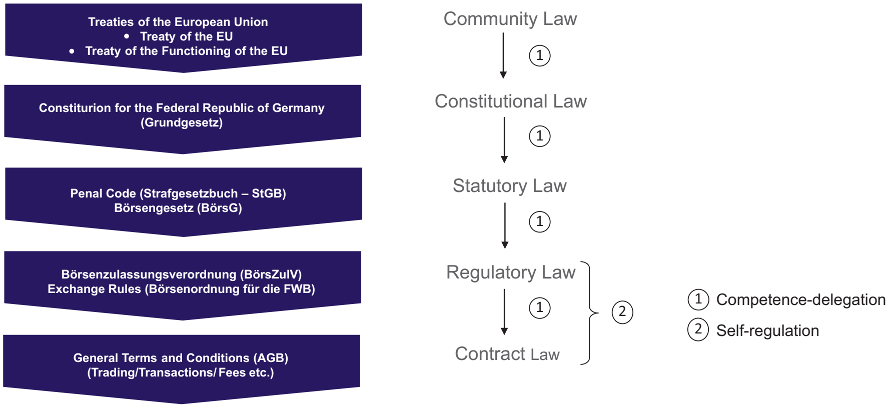


Fig. 2.6 Regulatory framework of Frankfurt Stock Exchange (Frankfurter Wertpapierbörse)

trading procedures). The overall objective of these exchange rules is the provision of a fair and orderly market, as well as investor and system protection. Thus, exchanges, in so far as they set, supervise and enforce their own rules, so-called **self-regulatory organizations** (SRO).

4. Application to local markets: Rules and regulations are specified by the exchange organization in consultation with market participants (the general assembly, board of directors, user groups, traders, back-office and IT staff, external experts, lawyers, market architects, etc.). The exchange's draft rules may be subject to approval by its general assembly or board of directors and the national or federal securities exchange commission, or other appropriate supervisory authority (or authorities in the case where supranational bodies are involved). Where government supervisory authorities are involved in the approval, they check the proposal in legal, market, and technology terms and decide to either approve or reject. If approved, the rules and regulations are implemented within a determined period of time. If rejected, they must be amended to include, change, or remove provisions as required.

2.3.3 Technology

Exchange organizations are high-tech organizations that enable their members to participate directly in the price discovery process. IT performance and reliability are a key success factor for any exchange organization worldwide. Essential for a fully fledged exchange platform is the concentration of liquidity in one single order book per product. The order book must be transparent, market access must be decentralized, and monitoring and surveillance must be centralized (Fig. 2.7).

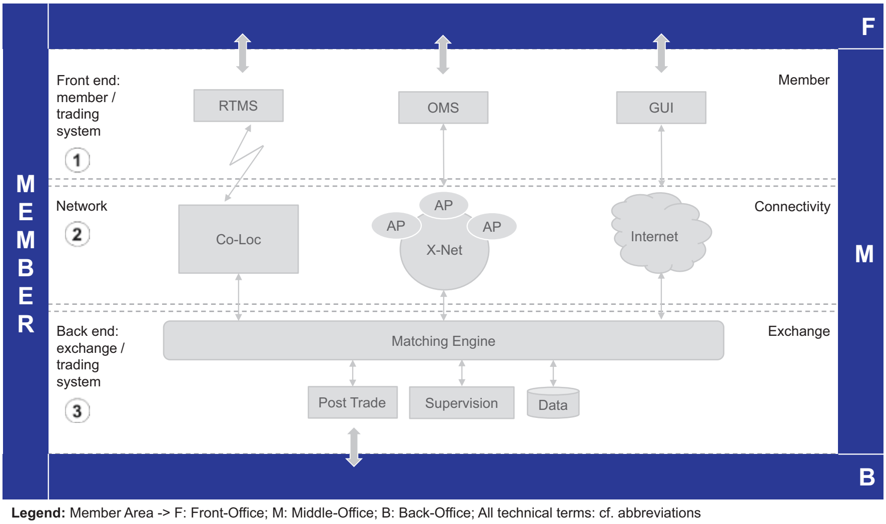


Fig. 2.7 Technological layer

The advantages of IT in the exchange business include the following:

- IT enables the exchange to go international, to integrate the value chain vertically and horizontally.
- IT allows members a decentralized access to price discovery around the globe.
- IT grants transparency and an effective and efficient market surveillance.
- IT assures better member support, more user-oriented functionality, and better overall risk management.
- IT can allow for extended trading hours (the three time zones).
- IT offers economies of scale and scope.

The functional elements of an exchange platform include the following:

1. The *member area*, where trades are generated either through a trading decision, an investment decision, or a decision made via a software component (e.g., program trading). In this area, the orders are handled by entering, deleting, changing, and holding orders. Also the position-keeping, basket-trading, and handling of conditional orders usually take place in this area. The necessary trading information is delivered directly from the exchange.
2. The *network* ensures constant connectivity that is managed by the respective exchange organization, sometimes in nanoseconds (latency matters!). There are three fundamental types of connectivity:

   (a) *Wide area networks*, sometimes covering all times zones: Access points enable members to locally access these networks, which are usually customized and tailor-made to fulfill the requirement of exchanges for high-volume data transfer as well as low latency, reliability, transparency, and—most important—fairness.

(b) The *Internet* is used to route orders.
(c) A *member's co-location* is installed as close as possible to the matching engine—this is the fastest way to be connected with price discovery.

(3) In the back end, actual price discovery takes place, driven by the central system of an exchange. It is the place where trading and trade management as order management are conducted as well as market control and market supervision.

The main elements of a back end are the following:

(a) *Trading*. Building up and maintaining the central order book is the main task. The trading module provides various market models, such as the order book and its related matching algorithms. This includes all of the functions that relate to the capture, processing, and execution of orders. The trading module also provides facilities for off-order book, bilateral trading (e.g., indications of interest and addressed offers), and trade and transaction reporting.

(b) *Trade management*. This module covers all of the post-trade facilities provided by an exchange. All trades (both those matched and reported) are handled. This module allows enquiry; the entry, limited modification, and deletion of trades; trade publication; and the management of trade reversals and other post-trade correction facilities. This module also passes trades on to the CCP and settlement organization for clearing and settlement.

(c) *Information dissemination and management*. This module immediately disseminates, from the exchange, all information that results from trading activity (recalculated indices, news, etc.) to the marketplace (members, surveillance personnel, data vendors, etc.). Most stock exchanges provide their own value-added information services that complement raw market data.

(d) *Market supervision and control*. This module includes monitoring and controlling the market on a day-to-day basis. It also provides features for handling exceptional situations. The main focus here is on the order book and trading activity.

(e) *Data and statistics*. This module maintains data, and it ensures statistical completeness, accuracy, and consistency of all data concerning members, issuers, and products. It also enables the production of a wide range of reports and statistics. And, most importantly, because this data are disseminated to members in real time, they can be used for real-time risk management and programed trading.

2.3.4 Stakeholders

A stakeholder is any party with an interest in the stock exchange as part of its internal and external environment. On the customer side the relevant stakeholders are:

- Issuers
- Intermediaries
- Investors

Relevant stakeholders on the regulatory side are:

- Regulators
- Administrators
- Surveillance authorities
- Legislators
- Politicians

And the relevant stakeholders on the governance side are:

- Shareholders through general assembly
- Exchange councils and other exchange-related committees

2.4 Customers of an Exchange

2.4.1 Issuers and Intermediaries

Critical success factors for *market participants* to trade on an exchange, and respective deliverables of the exchange operator and clearing house are as follows:

- Liquid and attractive product offerings: In the form of a diversified range of asset classes with a distribution network that is balanced between market makers and flow providers, and a competitive fee and pricing mechanism for all market participants.
- Reliable and high-performance IT infrastructure: Providing top performance in trading and networking as well as high stability, competitive service pricing, and a flexible architecture that also allows for short time-to-market product launches.
- Market integrity ensured by reliable rules and regulations: Covering all regular and extraordinary trading events—especially exchange rules mirroring international regulatory standards that allow for trading in different time zones.
- Mitigation of counterparty risk and efficient risk management: Guaranteed by the CCP that manages the fulfillment of obligations between participants and that ensures market integrity through the highest standards in real-time risk and default management.

Issuers create and deliver their product (the listed stock) by conducting an IPO. An issuer can have various reasons for a listing:

- Strengthening its equity basis for organic projects (big strategic projects) and inorganic growth (mergers and acquisition)
- Getting easier access to equity through capital increases
- Realizing capital gains and/or exit possibilities for existent shareholders
- Increasing public awareness: branding, image, etc.

In conducting an IPO, every issuer is pursuing a secondary market strategy centered around the stock price. This strategy may include additional elements like becoming part of an index or being represented in the derivatives markets. It may also be considered an option to get cheapest shareholder's equity. Regarding the secondary market, issuers evaluate marketplaces and/or segments based on criteria such as:

- Fast and inexpensive execution (high liquidity, low transaction costs, prompt and reliable settlement)
- High trading comfort (user-oriented functionality, good governance, attractive product range, reknown image/brand)
- Market integrity (fair pricing and regulation, optimized transparency and publicity, high investor protection, strong insider security, and effective surveillance)

The intermediary, being a member of an exchange, has to permanently fulfill certain criteria regulated by both law (Exchange Act) and the rules and regulations of the exchange. The following list summarizes the criteria that are exemplary for the Frankfurt Stock Exchange[9]:

- Guarantee a reliable management with at least one management member in possession of the professional qualification necessary for the exchange business.
- Ensure the orderly settlement of transactions concluded at the exchange.
- Provide evidence of equity capital in the amount of at least 50,000 EUR (such capital requirement does not apply for credit institutions and financial service institutions).
- Have the necessary economic capacity to participate in an orderly manner in exchange trading.
- Name at least one trader admitted to trading at the exchange.
- Settle its exchange trades via a CSD and provide an accounting connection to either Deutsche Bundesbank or any other central bank within the EU with direct connection to the EU payment system **TARGET2**.
- Clear its exchange trades via Eurex Clearing as CCP either as a direct (**general clearing member**) or an indirect (**clearing member**) member.

If the intermediaries are not the owners of an exchange, the exchange is called **demutualized**. Most of the existing big, regulated exchange organizations are partly or fully demutualized, and are themselves listed on a stock exchange.

Both issuers and intermediaries have specific contractual and legal relationships to an exchange, as first described in Sect. 1.8 and then, more thoroughly, in Chap. 8.

[9] Bullets 1–4: 19 Abs. 4 BörsG; bullets 5–7: §§ 12–14 BörsO FWB.

2.4.2 Investors: The Investment Process

In a stock market, an investor puts money to buy a stock to create more money when selling it, thereby taking risk to result in positive return, called profit or capital gain. An investor in a stock applies fundamental principles, e.g., the risk return principle and the principle of diversification. The investment process in the investor's view in general follows the structure shown in Fig. 2.8.

Professional investors, e.g., asset managers, are called the buy side. At the core of every asset manager's business is the investment process.[10] The investment process is a discipline that is still subject to continuous efforts for professionalization and standardization. The reasons are manifold. First, clients are becoming more demanding and the evaluation of the investment process is an integral part in their choice of the investment manager. Furthermore, in Europe an increasing number of institutional clients rely on consultants in their choice of asset manager, a development pioneered in the Anglo-Saxon investment business. These consultants also put the quality of the investment process at the top of their priority lists, when it comes to evaluating asset managers. And finally, asset managers themselves recognise the benefits of a rigorous investment process for their business.

The investment process seen as information flows is shown in Fig. 2.9.

First, information is created at the exchange through price discovery. Then this information is disseminated as market data by companies like Reuters and Bloomberg. Data as "raw material" are also used for new products like indices, which might then be used for structured products or derivatives, e.g., on an index.

Next step is data management, which feeds the software running investment solutions. The product of this software is then used either by asset advisors consulting asset owners or directly by the latter within the investment process.

2.5 The Elements of the Value Chain

A trade takes place through price discovery (1), is next cleared between the clearing house (2) and the intermediary, and is then settled between the settlement organization (CSD/ICSD: (3), including custody: (4)) and the intermediary or the (end)investor. This process defines the value chain (Fig. 2.10).

(1A) The **investor** in a specific stock is the flow creator. There are different types of investors:

- Retail investors: private persons

[10] The investment process is described in detail in Book II, Chapter XII.

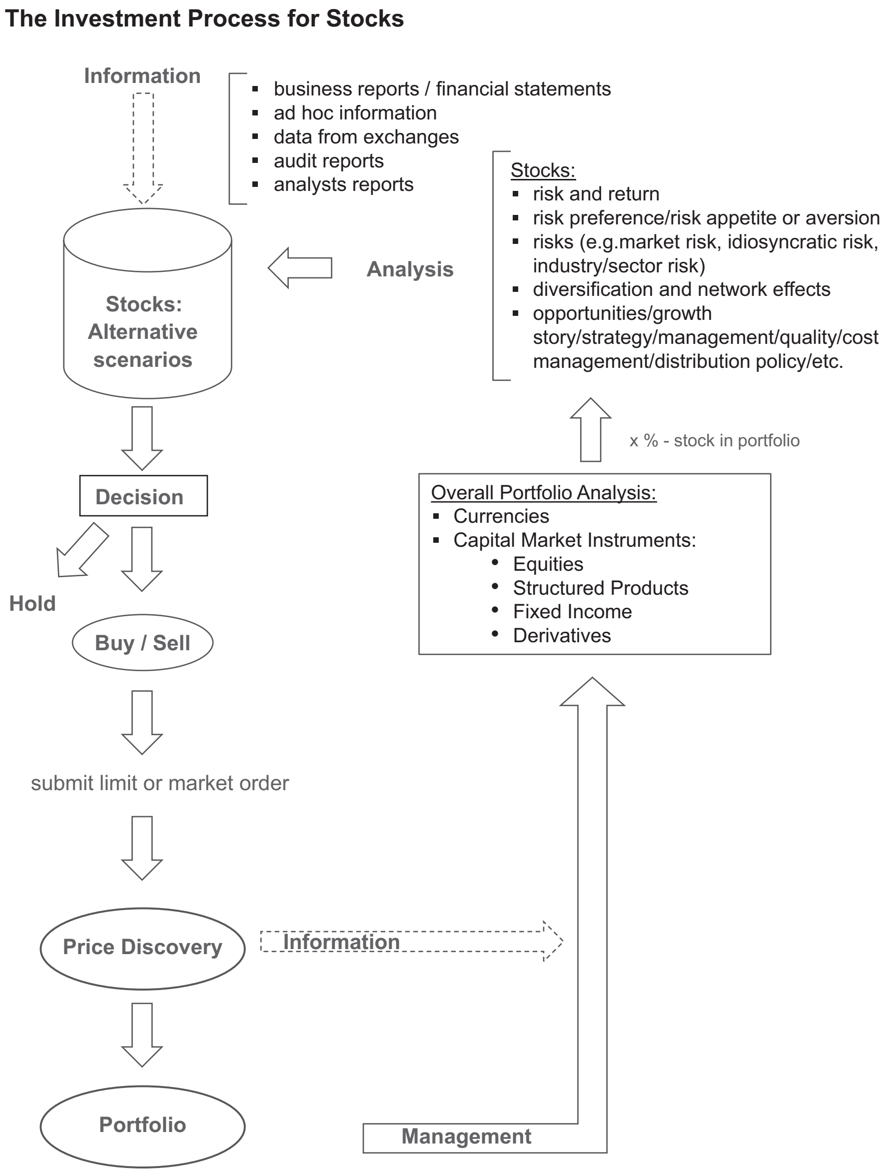


Fig. 2.8 The investment process for stocks

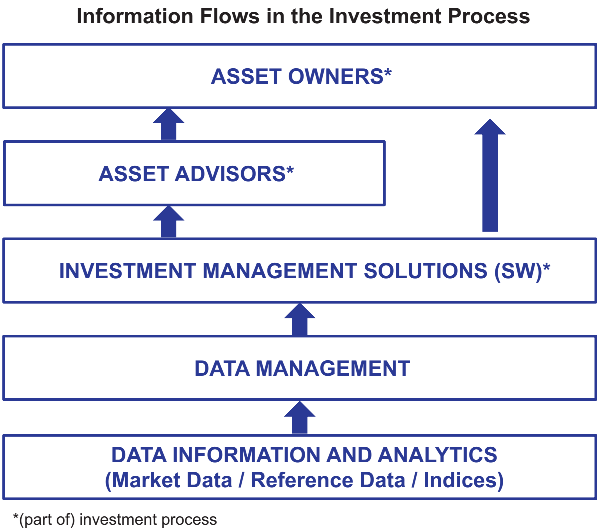


Fig. 2.9 Buy-side flow: data, software, and services

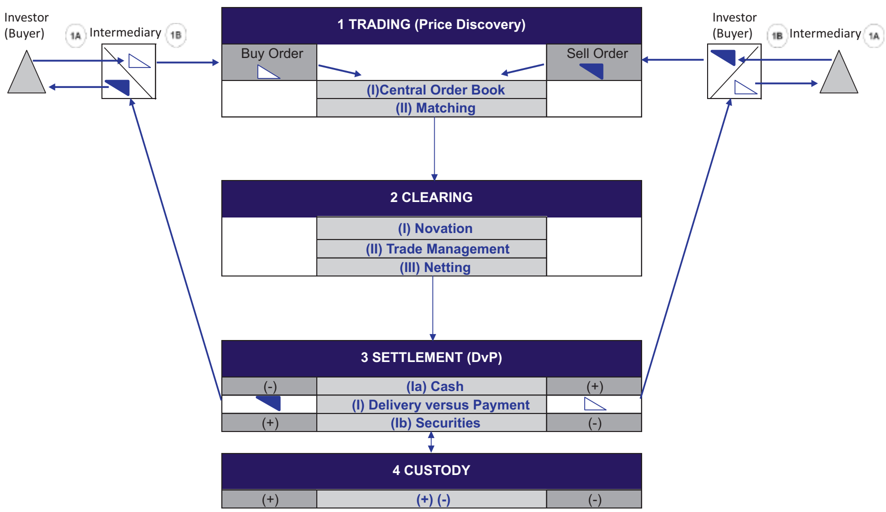


Fig. 2.10 The value chain: functionalities and flows

- Buy side: institutional investors (pension funds, insurance companies, hedge funds, etc.)
- Sell side: intermediaries (see 2) trading on their own account

(1B) The **intermediary** between an investor and a stock exchange is a bank or a broker. A bank as intermediary may act from its own account as a dealer or proprietary trader (intraday or long-term/strategic), or as a broker, simply routing the orders directly to the stock exchange. A broker does only order routing and therefore does not have a proprietary position in a trade.

The intermediary is a customer and a member of the stock exchange and therefore:

Regulated by securities and stock exchange laws

Surveyed by a stock exchange authority as well as respective organizational units (i.e., market supervision and market monitoring)

Controlled by banking supervision

Monitored by a clearing house

The functions of the core elements of the value chain are as follows:

(1) The **stock exchange** is the marketplace where bid and offer orders are collected and cumulated in a central order book, and where the price discovery takes place. In an order-driven market, this is done by predefined matching algorithms; in a quote-driven market, it is done by one or several market makers. Price discovery is at the heart of what an exchange does, and it requires that the market be fully transparent, monitored, and surveyed so as to create trust.

(2) The CCP or clearing house is legally the counterpart to the intermediaries, directly to the general clearing members (GCM) and indirectly (namely through the GCM) to the regular clearing members which are all trading participants (and sometimes also intermediaries for trading participants). The main tasks of CCPs are:

- Multilateral netting: to reduce transfer volume
- Risk management:

– Counterparty risk: ideally real time for all members through margins and other post-default backings
– Market risk: managed by monitoring *all positions* in a specific stock in *all stocks* listed at the exchange
– Ensuring through **margin calls** that all market participants fulfill their margin requirements promptly and permanently
– Protect the rest of the market in the event of default by (a) terminating the membership (e.g., Lehman Brothers) and (b) default management process

- System protection:

– Liquidity management

– Trade management: risk reports, tool, etc. (some of which might be on exchange side)

- Operability:

– STP and standard auction
– Processing of transactions
– Post-trade actions

(3) The CSD assures delivery versus payment of a trade in a stock for local markets (CSD) or cross border (ICSD). The securities accounts and the safekeeping of the shares are managed within a separate custody company and electronically interfaced with the CSD/ICSD.
(4) The **custodian** covers safekeeping and administration for the intermediaries. The administration includes accounting and also all necessary actions connected to a stock, like capital increase and dividend payment.

Value chain organizations, i.e., exchange, clearing house, CSD (or international CSD), and custodian, are the backbone of a financial system, be it on a local or even on a global level. Usually they operate fully electronically, and are also electronically interconnected. If they are fully integrated in one legal organization, they are said to be vertically integrated (i.e., Deutsche Börse Group, CME, ICE). If they are integrated with another organization of the same kind, they are said to be horizontally integrated (i.e., Euronext).

2.5.1 Trading: Price Discovery

The objective of price discovery is to match offer and demand of a specific stock in time, and in a consolidated order book, in order to find an execution price by applying specific rules (matching algorithms) (Fig. 2.11). The most common execution algorithm is the *maximum executable volume with lowest surplus* rule. When two different prices are possible, additional criteria and priorities have to be applied to get one execution price.[11]

A precondition to build this kind of market is fungibility of the product and, therefore, a standardization of the stock, usually specified by law. To build the central order book from scratch, the round lot (buy and sell side) has to be defined, before priorities are applied.

Price discovery can take place in different forms: order driven or through market making (Fig. 2.12). Order driven means that all incoming orders are matched against the spread or stored in the central order book, applying the price-time priority. Market making has two possibilities: single or multiple market making.

Price discovery takes place to open or close a market (usually through the opening/closing call) The same procedure is typically used to restart a market fol-

[11] See Chap. 4.

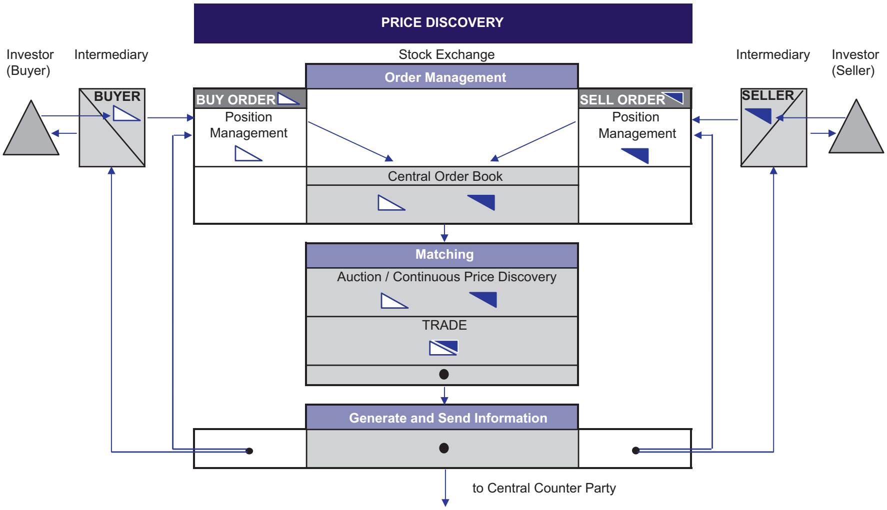


Fig. 2.11 Elements of the price discovery process

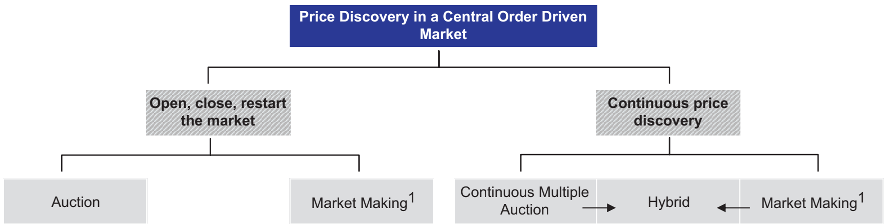


Fig. 2.12 Price discovery in a central order-driven market

lowing a trading halt. After the market is open, continuous price discovery within a defined period of time (trading time) takes place. The form may differ from multiple auctions only to market making or hybrids of the two forms, to continuous price discovery in just executing incoming orders against the spread in the central order book and if not possible store them, first according to price and within the same price according to time (Fig. 2.12).

Price discovery only takes place when the market is open. The market can be halted if one or more of the following criteria are not fulfilled:

- Orderly price discovery:
  - Both sides of the central order book have posted bids and offers.
  - The matching algorithm has to be consistent all the time.
- Equal access for all market participants to all key functions.
- Equal information for all market participants.
- Complete data integrity.

Scope, contents, and procedures of a fair and orderly price discovery are described through a market model.[12] The specification of a modular market model (MMM) includes:

- User groups and customers
- Trading form: call auction, order-driven market, market making, or a hybrid form including the matching algorithm
- Trading parameters: like transparency of the order book, different order types or the trading phases

Defining a tailor-made and customized market microstructure is an optimization procedure under the premises of creating the biggest possible liquidity pool in a specified segment. On the one hand, each module of the MMM interacts and interferes with the other modules, and on the other hand, the specification of a market model is elaborated sequentially from step I to step V.

To build the market model, in steps I and II, customer needs, user requirements, and particularities of the traded product determine the selection or choice. Then, in step III, the market model is defined either as a quote-driven or as an order-driven market. In step IV important trading parameters as shown in Fig. 2.13 have to be fixed.

The criteria of acceptability of a specified market model should—with reference to Thomas Kuhn's criteria of the acceptability of a theory[13]—be:

1. Consistency to guarantee fairness[14]
2. Agreement with market observations
3. Simplicity or the avoidance of unnecessary complexity especially concerning functionality
4. Breadth of scope to cover the chosen segments, e.g., products and customers
5. Fruitfulness, including conceptual integration and fertility for users, customers, products, and services

Furthermore, the market model has to fully reflect the rules and regulations of an exchange and is also fully mirrored in the software running the back end of the trading platform. The microstructure, defined through the MMM, has to be approved by the surveillance authorities and/or the regulators and implemented in the market, where it is eventually operated by the exchange.[15] The implemented market model must grant the overriding premises of investor protection and maximization of liquidity.

---

[12] Cf. [2], Equity Markets in Action, pg. 10ff.

[13] Cf. [3].

[14] Including equal treatment regarding access, transparency, information (ad hoc information and market data), fees, and functionality.

[15] The rollout of a new market model is a major technological project with key elements like member education and training (traders, IT departments, middle and back-office); tuning the system in terms of latency, volume, and functionality; getting started in time with members, surveillance, market operations, and other dependent entities like CCPs, settlement, and custody organizations.

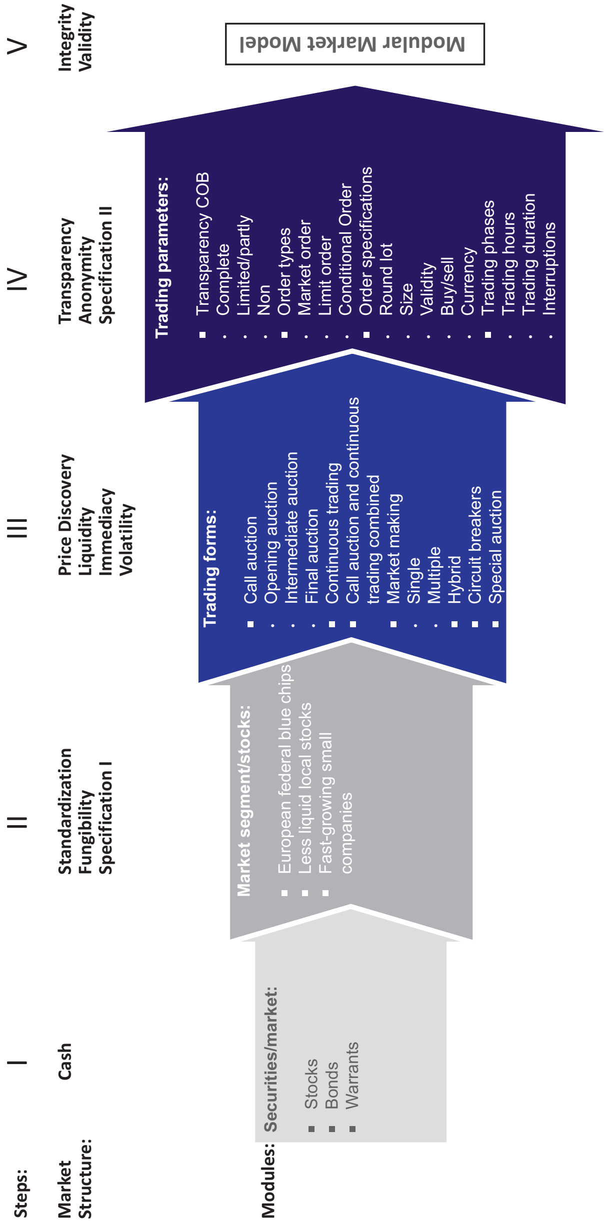


Fig. 2.13 The modular market model

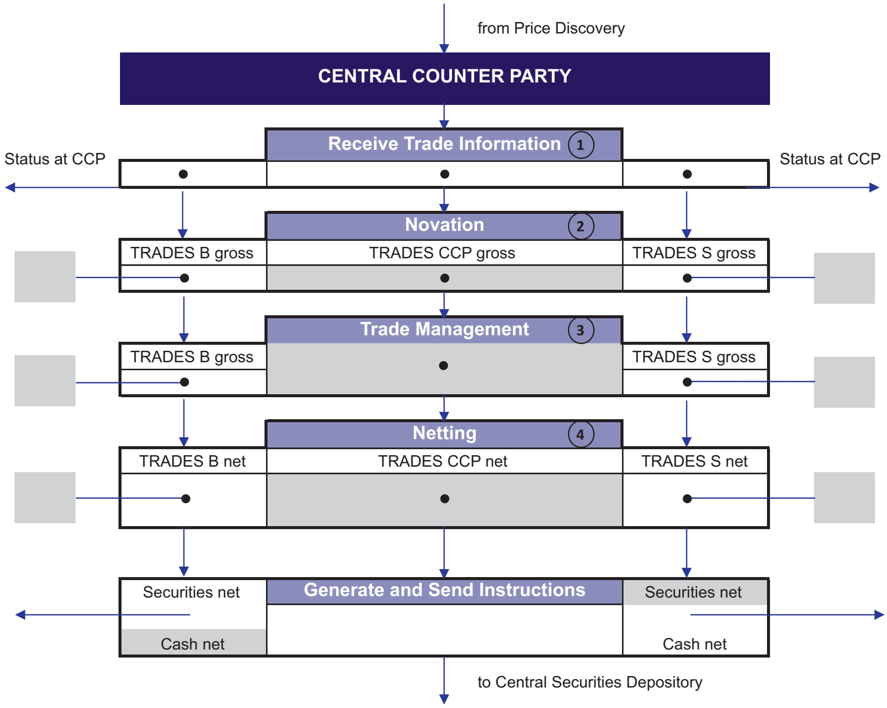


Fig. 2.14 Elements of a central counterparty

2.5.2 Clearing: Risk Management and Market Stability

Within the value chain, the CCP is the most important module to determine the risk profile of the industry. A CCP receives trade information from the exchange (1) to continue and to complete the transaction. Besides price and size, trade information contains specific master data like the symbol of the stock/security (e.g., DB1 for Deutsche Börse Group), the International Security Identification Number (ISIN), and, for local trades, some local master data (e.g., in Germany the so-called Wertpapierkennnummer (WKN)). The incoming information is used for novation (open offer) (2). Based on this process, trade management (3) and netting (4) take place consecutively (Fig. 2.14).

An intermediation between the trading activity and the settlement process takes place via a clearing house, a so-called CCP. There are several reasons for the development of such CCPs:

1. Reduction of gross risk exposure and enhancing capital efficiency through multilateral netting, therefore reducing counterparty risks and risks of trade.
2. Improving market structure: A CCP entering through novation (open offer) as counterparty in every trade guarantees specifying and monitoring of counterparty risk. A CCP takes over the counterparty risk of the market participant

directly (CM) and indirectly (through a GCM) and enforces strict risk control and adequate collateralization of open positions.
3. Assure anonymity.
4. Efficient use of collateral.
5. Neutral valuation of positions.
6. Improve risk management, payments, and delivery of stocks by IT on a real-time basis.
7. Improve market integrity by an early warning function and by avoiding a system collapse due to the failure of a market participant, e.g., Lehman (domino effect), in a timely manner through permanent margining.
8. Organize buy-ins and guarantee delivery.

The governance of a clearing house clearly contributes to its enhanced market integrity in preventing excessive risk taking by its members, simplifying market connectedness, and optimizing the collateralization of markets and counterparty risk management. The CCP is an integral part of a global strategy of exchange organizations.

A CCP that is registered with the regulator, is compliant with CPSS-IOSCO principles for financial market infrastructure, and is fulfilling a capital requirement of risk-weighted 2 % of their trade exposure is called a qualified CCP.

A CCP can cover local markets, regions, and even different time zones, and it can cover either a single product category or several categories. Consequently, there is a variety of different CCP types, and a clear distinction between a clearing house, a CCP, and a qualified CCP. To date, about 100 CCPs are registered globally, with the most prominent being CME Clearing, ICE Clear, LCH.Clearnet, and Eurex Clearing.[16] Every CCP reduces causes and magnitude of systemic risks, due to the fact that a CCP:

- Acts as an independent risk manager
- Creates transparency
- Has on its cash account only own money or central bank money
- Mitigates market member's counterparty risk
- (Most importantly) does not trade any financial instruments from its own account

2.5.2.1 Novation

Novation:

- Defines the (new) legal and counterparty risk structure in between the intermediary, the stock exchange member, and the clearing house (Fig. 2.15)
- Changes the entire risk and liability structure of a marketplace because, on the counter side of a trade, the individual counterparty risk is replaced with the higher creditworthiness of the CCP
- Is a necessary precondition for netting

[16] C.f.: DBAG White Paper 2014, pg. 8.

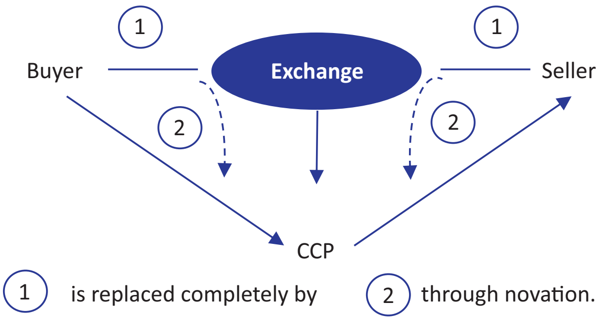


Fig. 2.15 Novation

- Is the basis for handling risk management outside the clearing house, starting with netting and continuing with the margining procedure

Legally, novation means that one obligation is replaced (1) by a new obligation (2), whereby the content of the new obligation (2) is identical with the content of the old obligation (1). The old obligation (1) is then no longer in existence and is to be treated as if it never existed.

What this means in terms of market and liability structure is shown in Fig. 2.16.

2.5.2.2 Netting

After novation, the netting of all buy and sell orders of a participant in a predefined time period and in a predefined product range (cross-products) takes place, usually once per trading day. Due to the netting, only the residual amount of securities and/or money is due. Through this netting procedure:

- Complexity and thereby costs for clearing members are significantly reduced.
- Volume of transferable securities and money is significantly reduced.
- Risk management is improved.

The effect of netting becomes obvious when comparing (for the same situation in terms of market participants and order flows) the three scenarios: no clearing, with bilateral clearing, and with multilateral clearing. The nature of the netting effect can be seen as quantitative or qualitative.

To start with the quantitative netting effects, the key performance indicators (KPI) are shown in the chart below.

Figure 2.17 shows a simplification of the market situation. Here the market consists of three participants or members (A/B/C) and six different financial order instructions, in this case one stock. In this market the netting effect is a factor of 10, comparing a situation without netting with a situation in which multilateral netting through a CCP takes place. The effectivity of netting itself increases with the clearing volume. The netting volume is a product of the number of market participants,

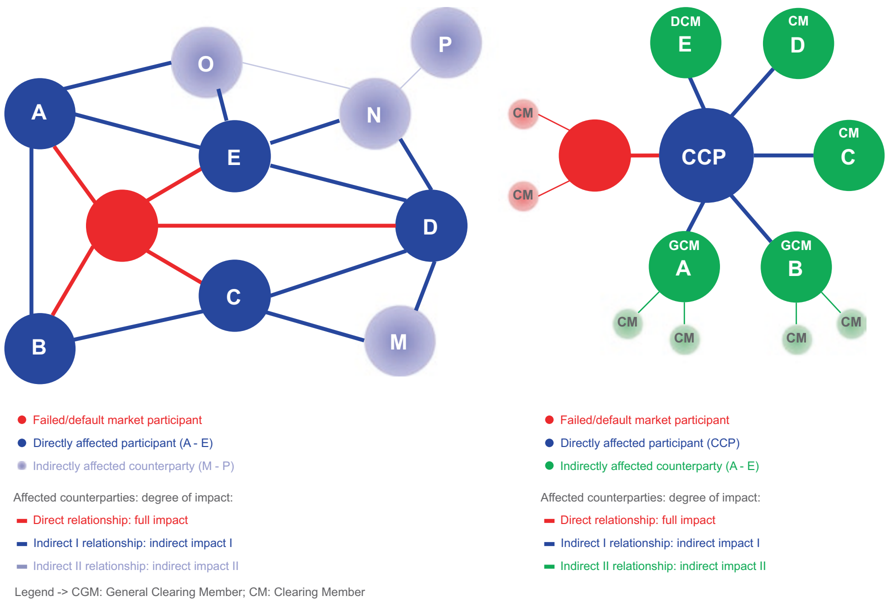


Fig. 2.16 Market and liability structure in bilateral vs. multilateral netting

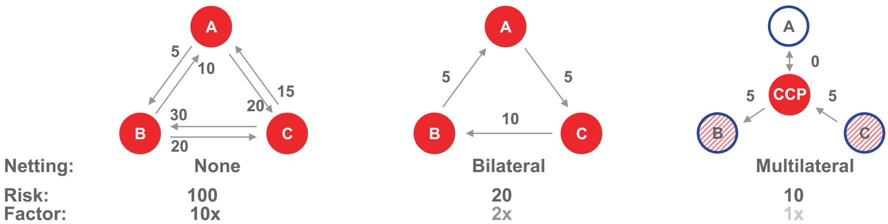


Fig. 2.17 The quantitative netting effects

the number of financial instruments traded, and the volume of market participant's open interest: The higher the number and the higher the volume, the bigger the netting effect and the cost reduction.

The qualitative dimension of netting is summarized in Table 2.1.

Hence the qualitative effect of netting to market integrity is the reduction of complexity and of financial and counterparty risk (Fig. 2.18).

Coming from no clearing to a clearing house the following is the case for exchange and clearing members:

- Mistakes in handling and administrating stocks go down.
- If this process is driven electronically, data and information handling is much faster.

Table 2.1 Qualitative dimension of netting

	Bilateral netting	Multilateral netting through novation
Market structure	• Very complex relationships: Everyone trades with everyone directly/bilaterally	• Simple: one central counterparty
	• Many dependencies	• Every trading party has one (and only one!) bilateral relationship to the CCP
	• Bilateral relationships are not known: intransparent	• Business relationships are transparent in the CCP
Liability structure	• All market participants are directly or indirectly affected	• Only CCP directly affected[a]
	• Danger of domino effect	• CCP is hedged against default of one of its members (cf. default management)
	• No information about magnitude of market impact (who, how much, etc.) due to opaque situation	
Liquidity structure	• Distressed market participants need liquidity; many counterparties	• Distressed market needs liquidity: one access partner (e.g., EUREX: Bundesbank $\gg$ collateralized access to central bank money)

[a]Depending on clearing structure, in specific cases (if a GCM fails) other clearing members might be directly affected as well

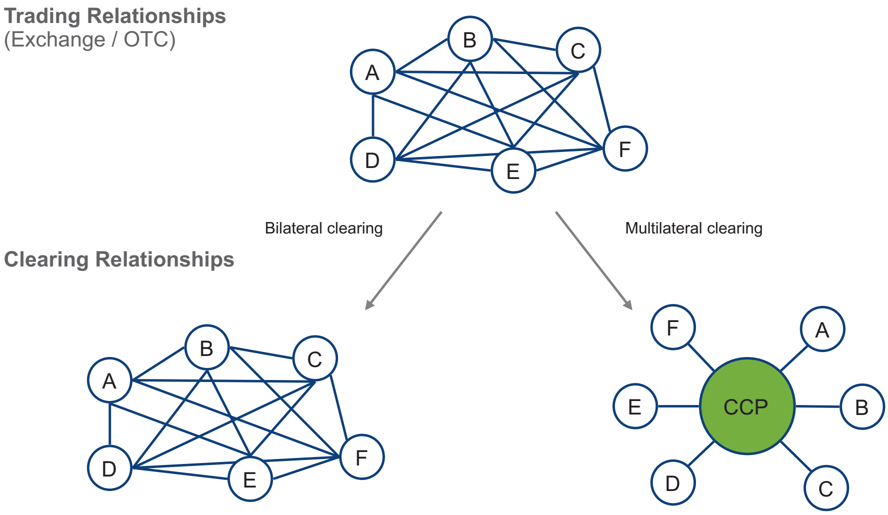


Fig. 2.18 Complexity of bilateral and multilateral clearing

Therefore:

- Costs go down due to reduced complexity and lower volume.
- Risk structure and risk management improve.

For the whole market structure, this means that:

- Transparency is enhanced.
- Surveillance/control improves.
- Liability structure with a CCP stabilizes a marketplace because defaults can be mitigated by a CCP.
- Arbitrage is reduced.
- Revenue potentials might be reduced because the CCP must be paid for services provided.

2.5.2.3 Risk Management

In this scenario the main risks of a CCP are:

I Short term:

- Counterparty credit risk: The default of a member, thus the credit risk, is the main risk for a CCP (market/liquidity risks are contingent on this credit risk of members). However, this risk is largely mitigated with the conservative margin that assumes a probability of default for members of 100 %, as well as the robust **lines of defense**. So this risk accounts for roughly one-third of a CCP's total risk. (Covered probability of default for members is a conservative 100 %.)
- Operational risk: This is mainly driven by service deficiency. Legal risk, damage to physical assets, and unavailability of services and systems. Due to the central role of a CCP these risks account for roughly two-thirds of a CCP's total risk.

II Mid- to long term:

- Business risk: for example, regulatory and macroeconomic developments
- Project risk: for example, the implementation of new functionalities or hardware

The risk profile of a fully fledged CCP is made low by margin, etc. and handled through stringent processes and product collateralization, especially also for the positions of members in default. Regarding the risk profile of a CCP, it is important to bear in mind that:

- CCP positions are always balanced
- CCPs are not trading on their own account:
  - No market exposure
  - No speculation
- Members default:
  - Special procedure and strong decision power of a CCP
  - Potential losses covered by margining and buffer

- Robust lines of defense[17]

The risk management of a CCP starts with member admission and continues with member monitoring. The applicant has to fulfill the following requirements[18]:

Apart from being licensed by the local authority to provide credit to clearing customers and to receive collateral in the form of cash and securities, an applicant must be under surveillance of a responsible local authority according to standards equivalent to the applicable regulatory standards in the EU, and the applicant is required to sign relevant clearing conditions. The clearing function contribution (which is linked to the risk exposure or at least to the minimum contribution amount) can be submitted in the form of collateral in securities and/or cash. Based on the asset classes that they are trading, members must fulfill minimum requirements to get the corresponding clearing license: the applicant needs to have (or to open) an account at the corresponding central bank (cash account) and CSD (securities and cash account), and also needs to provide evidence of a securities account and a pledged securities account with the central custodian. Additionally, various technical and operational requirements have to be fulfilled permanently.

A CCP evaluates, controls, monitors, and handles not only in stocks but, in the best case across all asset classes, different risks centrally, real time, and online. The most important risk is the counterparty risk:

The counterparty risk of all the general clearing members (GCM), and the regular clearing members (CM), is monitored by the permanent calculation of the margin requirements (Fig. 2.19). Ideally, this calculation (valuation) is done in real time, as for example at Eurex Clearing. Members have to cover their risk with collateral against a haircut. If the risk of a member is not covered, an intraday margin call will take place. Schematically this is explained in the following exhibit:

As a concrete example of the Eurex clearing house, on a monthly basis, one can start to deduct from the overall clearing volume of €16,343 billion the netted volume of €16,304 billion (Fig. 2.20). In case of the Eurex-CCP, a buffer of approximately 20 % is calculated. This enhances the resiliency of the system. It is like a security premium for the market, which amounts in this case to €12 bn. Adding the net margining requirement and the buffer results in the sum of the collateral needed to cover the total margin requirement. Eurex calculates this margin cross-product and in real time. Therefore, their data is also an excellent basis for the risk management of the clearing members, e.g., the banks.[19]

[17] With this structure, CCPs' lines of defense withstand an equity market drop of 30 % ([4], p. 23).

[18] For example: [5], Clearing Conditions for Eurex Clearing AG, Chapter I, Part 1, Top 2.1: https://www.eurexclearing.com/blob/clearing-en/51612-136778/238376/34/data/clearing_conditions_en_ab_18_11_2013.pdf.pdf.

[19] For example: [5], Clearing Conditions for Eurex Clearing AG, Chapter I, Part 1, Top 3: https://www.eurexclearing.com/blob/clearing-en/51612-136778/238376/34/data/clearing_conditions_en_ab_18_11_2013.pdf.pdf.

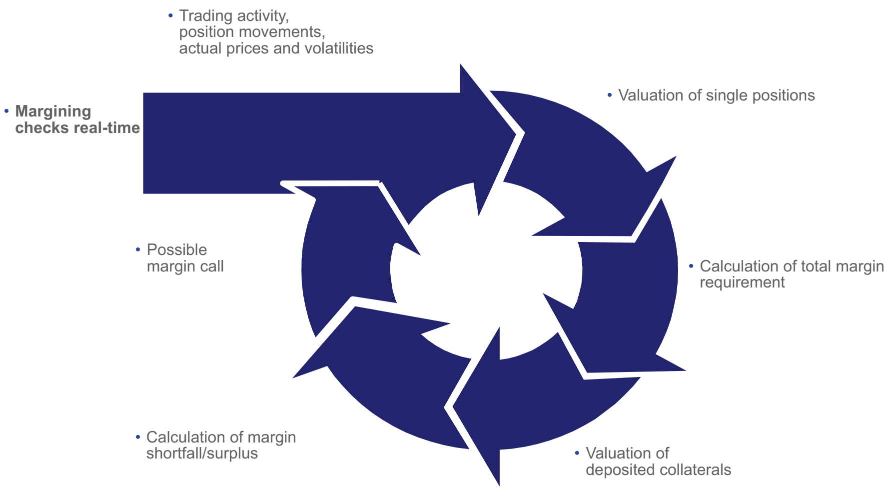


Fig. 2.19 Permanent margining of counterparty risk

Regarding the systemic risk of a capital market, the CCP has to be considered in a broader environment. In the chart above, the following phases of the risk management process and respective effects of a CCP can be distinguished:

1. Phase I: Elimination of €16,304 billion by multilateral netting; the systemic risk is reduced by this amount.
2. Phase II: To secure the counterparty risk of the clearing house, the whole amount has to be managed as follows:

   (a) Parameterized haircuts for each individual collateral
   (b) Cross-asset class collateralization
   (c) Real-time *portfolio management*
   (d) Timely margin calls when necessary

With this ongoing margin management, the counterparty risk of a CCP is permanently covered with a buffer of approximately 20 %.

The overview provided in Fig. 2.20 outlines that the CCP, in acting like a hub, is *the* central part that gives an indication of what the market risk and the systemic risk could be. If the respective product is of global importance, risk management can be monitored even on a global level (Fig. 2.21).

2.5.2.4 Lines of Defense

If a clearing member cannot fulfill the margin requirement (net position plus buffer), a margin call is made. To cover the position, cash and securities with a corresponding haircut can be brought in. A haircut can be delivered in:

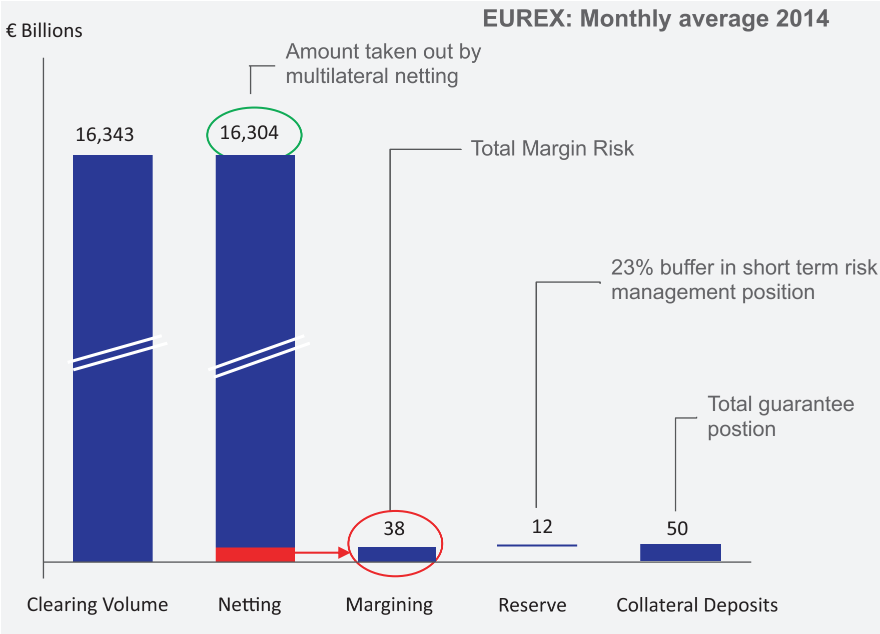


Fig. 2.20 Risk management and mitigation: the systemic risk

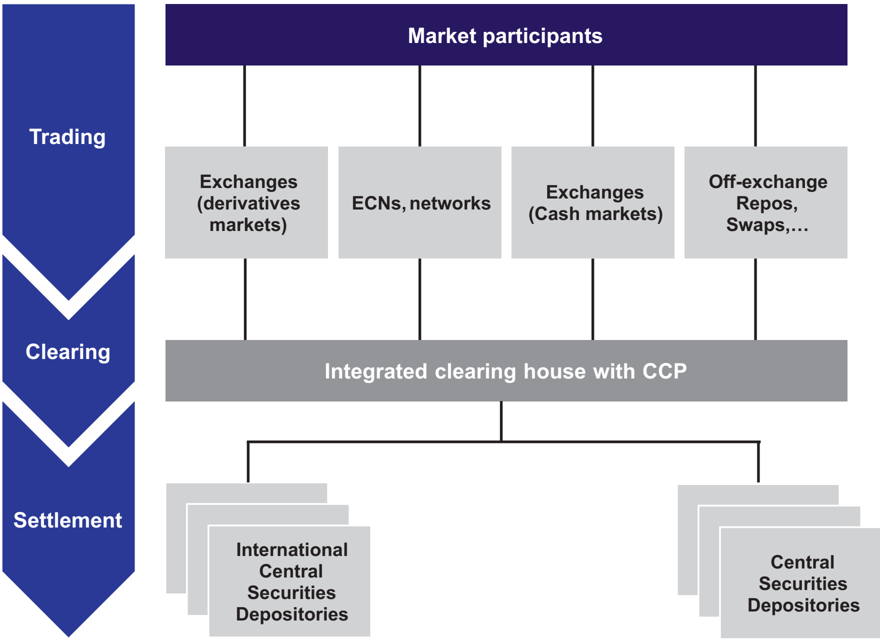


Fig. 2.21 CCP and enlarged value chain

- Cash (no haircut)
- Fixed income (haircut based on rating of debtor)
- Equities (haircut starting at 30 % of the mark-to-market value for the best rated listed companies)

Systemically the lines of defense look as follows:

The important elements of this *waterfall* structure, like the Eurex Clearing's Lines of Defense that correspond to EMIR, are funded by the members. In monitoring and maintaining these lines of defense, a clearing house—if rightfully handled—cannot go out of business. On the contrary, taking out market distressed members in a timely fashion (e.g., Lehman Brothers), a CCP enhances system stability, thereby being part of the solution, not part of a problem. In general, this means that CCPs reduce the risks of domino effects in a crisis situation; they do so by replacing complex bilateral relationships with high interconnectedness (*spider web*) with a 1:1 relationship to a CCP. With their primary focus on risk mitigation and control, the CCPs act like shock absorbers for the financial markets (Fig. 2.22).

2.5.2.5 Default Management

In case of the default of one or several clearing members, a clearing house must protect its customers and minimize impact on a client's positions. A corresponding default management process facilitates the liquidation of the defaulter's portfolio (Fig. 2.23).

1. Client transfer and preliminary measures: A default management committee assists the clearing house in the default management process while the defaulting client's positions and collateral are transferred for hedging and auctioning.
2. Hedging: The hedging process reduces risks for the clearing house and stabilizes the portfolio for the following auctioning process.
3. Independent sale: The default management process includes the possibility for an independent sale of a certain liquidation group to another clearing member.
4. Auctioning: Establishing a fair price for the respective portfolio; this is the integral responsibility of the clearing house in the auctioning phase. Clearing members who are active in the liquidation group are required to participate in the auctions. Other customers may participate so long as they are compliant with the clearing member's bidding obligation.
5. Asset class resolution: Losses from possible remaining positions are charged to non-bidders in cases where they do not agree to enter into a residual settlement (that is, taking on these positions from the clearing house at auction prices or at the last mark-to-market price). If no member participated in the auction, the corresponding transactions may by terminated by the clearing house.

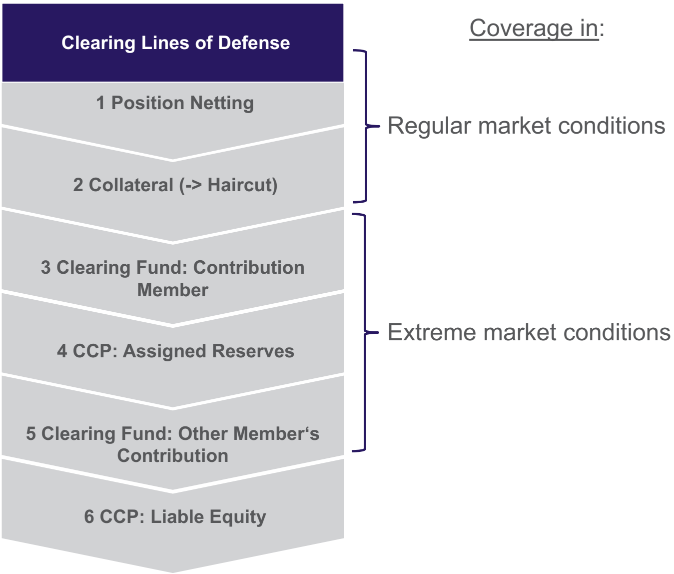


Fig. 2.22 Lines of defense

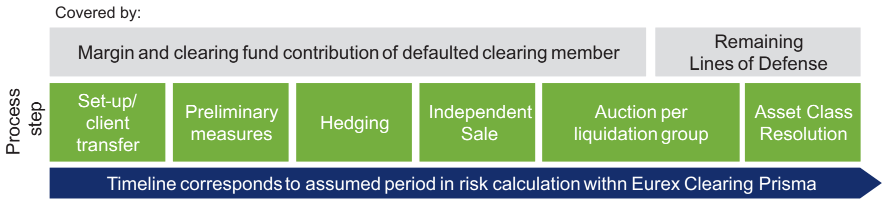


Fig. 2.23 Overview of the default management process framework of Eurex Clearing

2.5.2.6 Key Criteria for CCP Rating

Because they centralize counterparty credit risk, CCPs are perceived (especially by regulators) as relevant for systemic risk. The following main criteria are indicators for the stability of a CCP:

1. *Lines of Defense*: CCPs must provide a comprehensive risk management that contains sufficient collateral to cover losses (*defaulter pay* model), a mutual clearing fund, which covers extreme tail risk in case of member failure, and limited liability that provides certainty for members about their credit exposure.
2. *Access to Liquidity*: Sufficient liquidity is crucial for a CCP to cover obligations at any time. The liquidity source needs to maintain its reliability even under stressed market conditions (access to central bank liquidity is considered to be the most reliable source).

3. *Governance*: Structure of the CCP's management and risk committee must comply with regulatory requirements.
4. *Regulatory compliance*: The risk management framework has to reflect recent regulatory standards.

2.5.3 Settlement: Delivery Versus Payment

Settlement is the delivery of the traded securities against the payment in cash for the traded securities through a CSD within a specified time (usually $T+2$ working days) and stands for the completion of the trade transaction. In short: DvP. The overriding principle is delivery of securities (by a CSD) versus payment of cash (by central bank payment systems or a cash platform, i.e., in the EU: Target2) within a predefined time period (Fig. 2.24).

If the settlement takes place in a domestic market for domestic stocks, a CSD executes this transaction. If the transaction is cross border, an ICSD executes. Starting the settlement-process, the CCP determines the CSD or ICSD.

The settlement information and instruction to the CSD/ICSD come from the CCP and the stock exchange. All the information and instructions are checked by the settlement supervision. In the validation phase, the correctness and completeness of the settlement instructions are checked, and instructions may be changed or

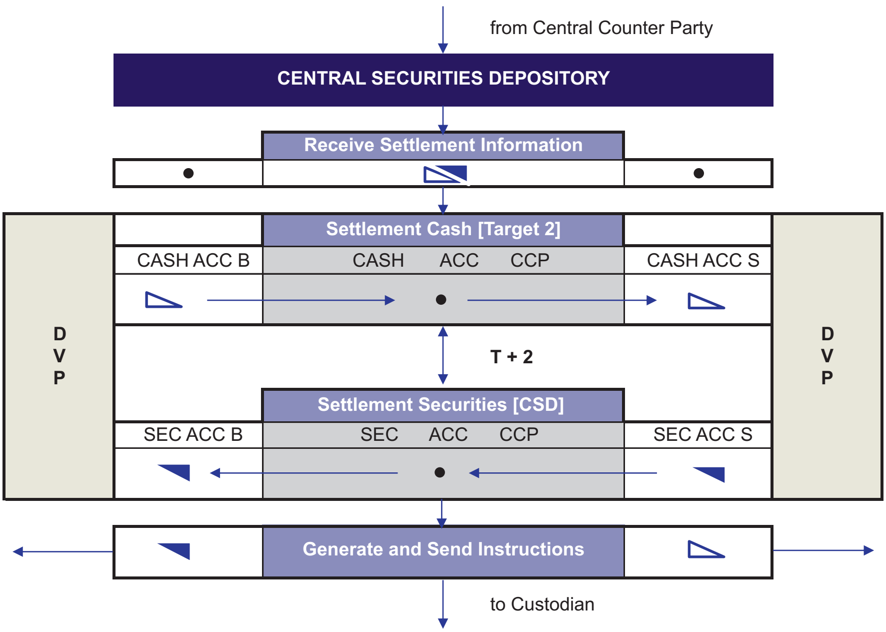


Fig. 2.24 Elements of a central securities depository

deleted. Blocking and release of instructions to matching immediately or on customers request are also part of the validation. Matching in settlement means bring together corresponding instructions from the two counterparties of a trade for the purpose of settling (thereby applying settlement-matching-rules). Based on the action, the status of the settlement instructions has to be set to *matched*, *unmatched,* or *advisory*.

For settlement (DvP), the balances of the securities (delivery) and the cash account (payment) have to be checked and, in case of insufficient need, to be corrected. Then the sequence of bookings can be optimized and eventually the debit and credit on cash and securities accounts takes place. Based on that, *settled* can be given as an information via the CCP and/or exchange to the customer.

There are two types of CSDs as shown in Table 2.2:

- CSD: Covers national securities (stocks, warrants, corporate and government bonds) in local currency and settles nationally.
- International central securities depository (ICSD): Covers international securities in numerous currencies and settles cross border.

If a CSD provides or delivers securities as security collateral to cover financing and credit facilities, this business is called a **pledge**. In order to enhance efficiency and to reduce complexity to make cross-border transactions in Europe faster and cheaper, *Link Up Markets* led by *Clearstream* was set up. On the one hand, (local) CSDs stay unchanged in function and mission; on the other hand, with Link Up, they can integrate cross-border functionality. This is a quantum leap in Europe's settlement area (Fig. 2.25).

The Link Up Markets model will replace the current inefficient setup with the following advantages:

- Single point of access for CSDs to participating markets
- Easy implementation of enhanced CSD links in central bank money leveraging TARGET2-Cash
- Reuse of efficient local infrastructure
- Absorbing differences in market standards
- Best-in-class CSD services for all asset classes (excluding derivatives) and multiple currencies
- Continuous harmonization of market practices
- Flexible extension of market coverage in Europe and beyond

Table 2.2 CDS vs. ICSD

	CSD	ICSD
Currency	Domestic	FX and domestic (multicurrency)
Securities	Domestic–domestic (intramarket/ national)	Cross border: foreign domestic (intermarkets/ international)
Custody	National	National and international

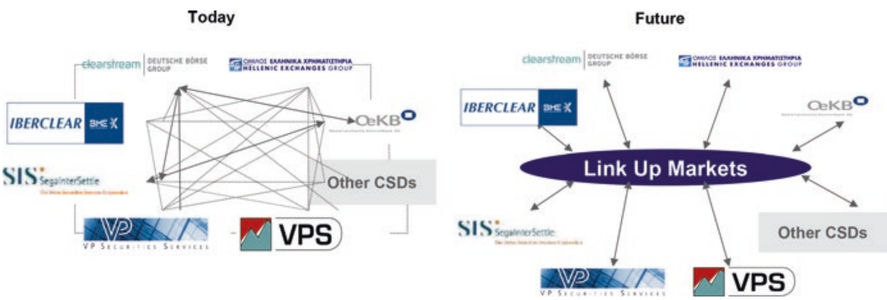


Fig. 2.25 Current and future European CSD landscape

	Europe	United States	
Equities	**Central Securities Depositories (CSDs)** ▪ Clearstream Banking AG Frankfurt ▪ Euroclear Belgium (CIK) ▪ Euroclear France (Sicovam) ▪ Euroclear Netherlands (Necigef) ▪ Euroclear UK (CrestCo) ▪ 17 CSDs in the Eurozone alone	**DTCC** ▪ User-owned CSD ▪ For equities and non Fedwire eligible bonds ▪ Without banking services	Equities
Fixed Income	**International Central Securities Depositories (ICSDs)** ▪ Services provided by Clearstream Banking S.A. Luxembourg and Euroclear Bank ▪ Securities depositories and Settlement agents for international debt and other instruments ▪ Banking license, with value-added services	**Bank of New York / JPMorgan Chase** ▪ Settlement agents for primary dealers in Fedwire eligible securities ▪ Banking license, with value-added services ▪ Utilize Federal Reserve securities wire assuring finality of settlement and payments	Fixed Income

Fig. 2.26 European versus US post-trade landscape

With respect to the post trade landscape, there is a significant difference in the European versus the US marketplace (Fig. 2.26):

2.5.4 *Custody: Administration and Safekeeping*

Custody means administrating and safekeeping securities for others. This enables settlement in exchanging securities between seller and buyer. A custody account, the equivalent to the money account, is established for each customer. Customers of a custodian are usually banks, not private persons or individuals. The account information includes details of the types, nominal values, and quantities or volumes of the securities held, as well as the name and address of the account holder (Fig. 2.27).

Administration delivers services in the following areas of a security:

- Clarifying and conforming ownership and all rights associated with it (dividend, voting, liquidation)
- Capital elements (e.g., dividends, capital increase for a stock)
- Tax

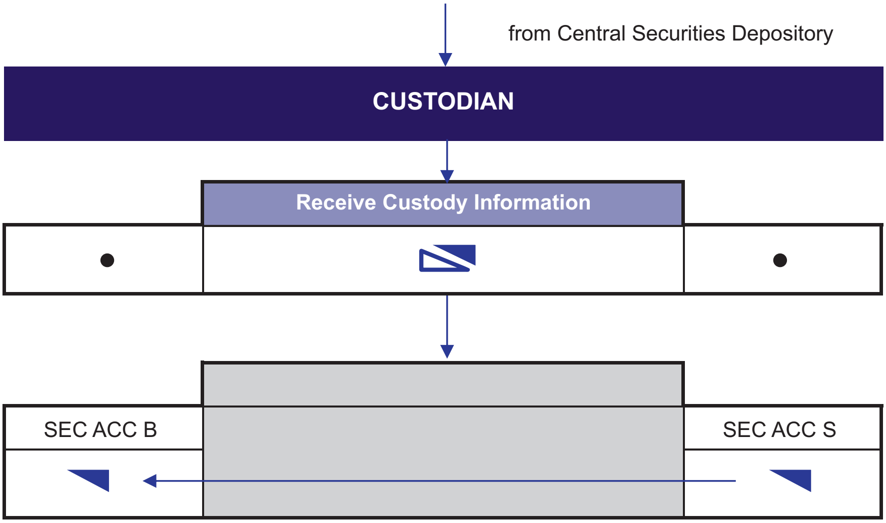


Fig. 2.27 Elements of a custodian

- Shareholder meetings (e.g., information concerning the voting rights, invitation, substitution)

These activities are summarized in Fig. 2.28.

The custody market in Europe is considerable, namely €10.2 trillion. Its structure is fragmented into the following segments (in trillion Euro)[20]:

- Market-secured funding: 5.1 (50 %)
- Central bank-secured funding: 4.5 (44 %)
- Trading with CCPs: 0.3 (3 %)
- OTC derivatives margining: 0.2 (1.5 %)
- Settlement: 0.2 (1.5 %)

A study carried out by Clearstream and Accenture revealed that global institutions manage their collateral or cash on an individual trading desk basis without any coordination, leaving a single institution with a number of discrete liquidity pools. This leads to expensive collateral use. The situation externally is similar: banks tend to have positions across a wide number of markets, and they maintain a discrete collateral pool in each.

Accenture found that the greatest benefits from overcoming these inefficiencies were reaped from enabling institutions to maximize liquidity, reduce financing costs, and lengthen the funding term. The study reported that, if these issues were resolved, the potential value from optimizing collateral so as to replace unsecured funding with a secured equivalent could be €3.8 billion. Additionally, the cost of maintaining excessive levels of collateralization with multiple settlement agents, involving greater legal costs and development of multiple interfaces with a variety of external providers and internal pools, could be around €400 million.

[20] See [6].

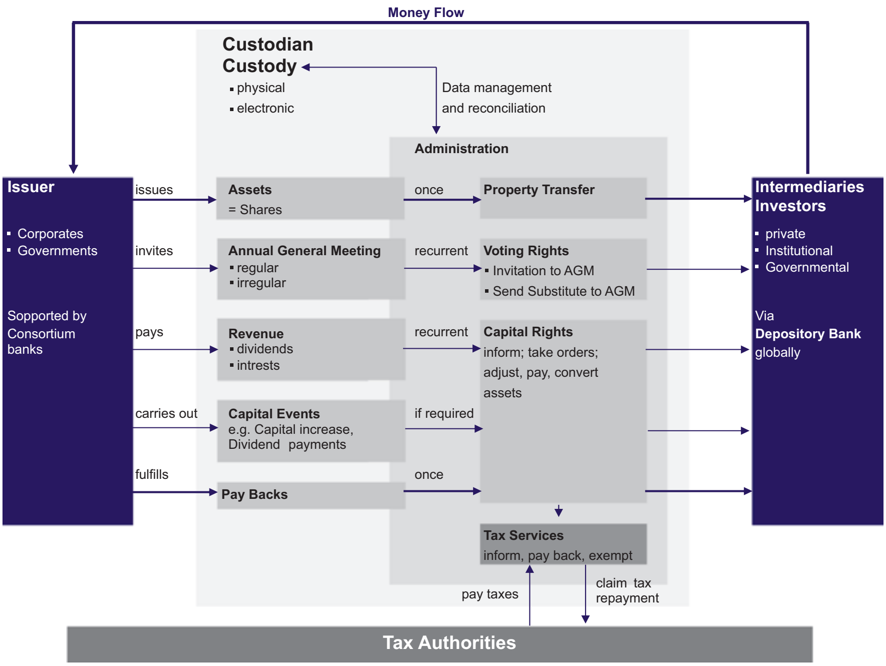


Fig. 2.28 Custodian—elements and flows

During the credit crisis, tri-party agent services in particular experienced a boom, as they allowed banks to better manage and cover their exposures through a neutral infrastructure provider. CSDs are particularly well placed to help the market overcome possible collateral shortfalls by optimizing collateral pools. Clearstream, for example, has long-standing experience in providing collateral management and securities lending services via its global liquidity hub. It can leverage this knowledge to help partners overcome collateral fragmentation and the related cost.[21]

2.6 The Objective of the Value Chain: Straight-Through Processing

If a trade is matched immediately and then directly cleared and settled, the process from starting the order at an intermediary to the receipt of the settlement confirmation at the same intermediary is called STP (cf. Fig. 2.29). The instant booking of cash for securities means processing within seconds or faster with one objective:

[21] See [6].

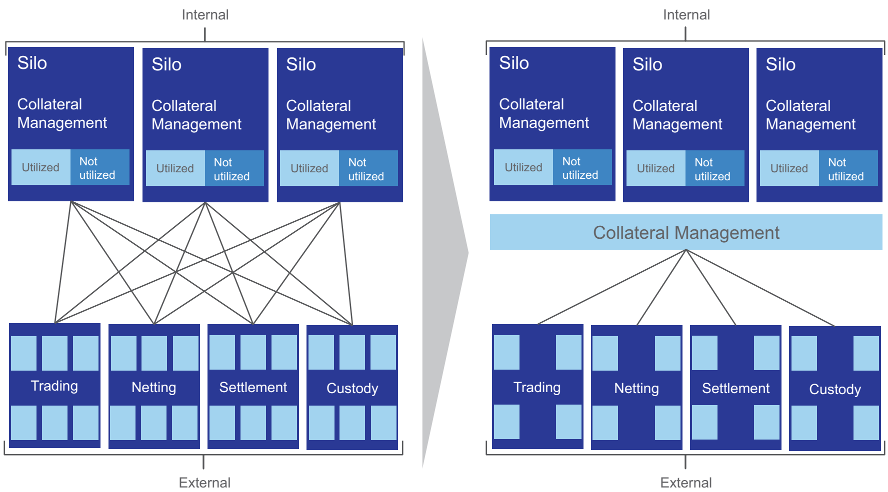


Fig. 2.29 Custodian: reducing internal and external fragmentation (Accenture 2014)

minimize execution costs and all kinds of risk.[22] An integrated value chain (such as Deutsche Börse Group has) generally offers instant clearing and following settlement: the usual round-trip time for orders ranges within microsecond trading at peak times.[23]

STP involves the circulation of money and securities: The role of the clearing house in this process is twofold: on the one hand, the CCP handles the exchange; on the other hand, the CCP assures the settlement (DvP) via its own balance sheets and in real time.

The security and the money flow within the value chain are outlined schematically in Fig. 2.30. The STP process in the securities' value chain ensures the execution of all processing steps via electronic media. Manual interference is not required. Data once captured cannot get lost, are protected from errors in manual processing, and only need to be entered once. That even applies when processing across the boundaries of separate asymmetrical units or even companies and institutions. A prerequisite for STP processing is a high degree of standardization along all steps, as well as reliable and stable points of execution of these steps (Fig. 2.31).

[22] For example volatility risk, counterparty risk, market risk, country and currency risk, liquidity risk, hedging risk.

[23] Round-trip time refers to the time order processing takes from an intermediary via order book and matching to the CCP and back to the intermediary.

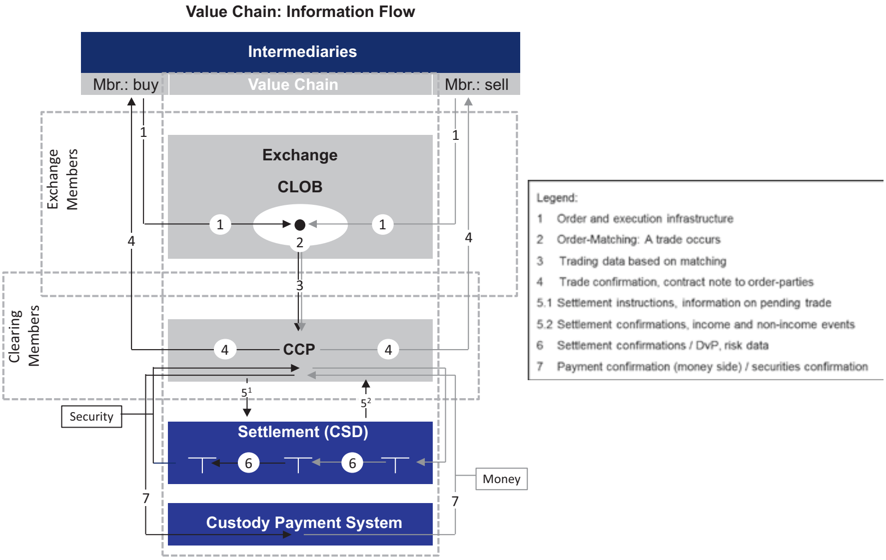


Fig. 2.30 Flow of information within the value chain

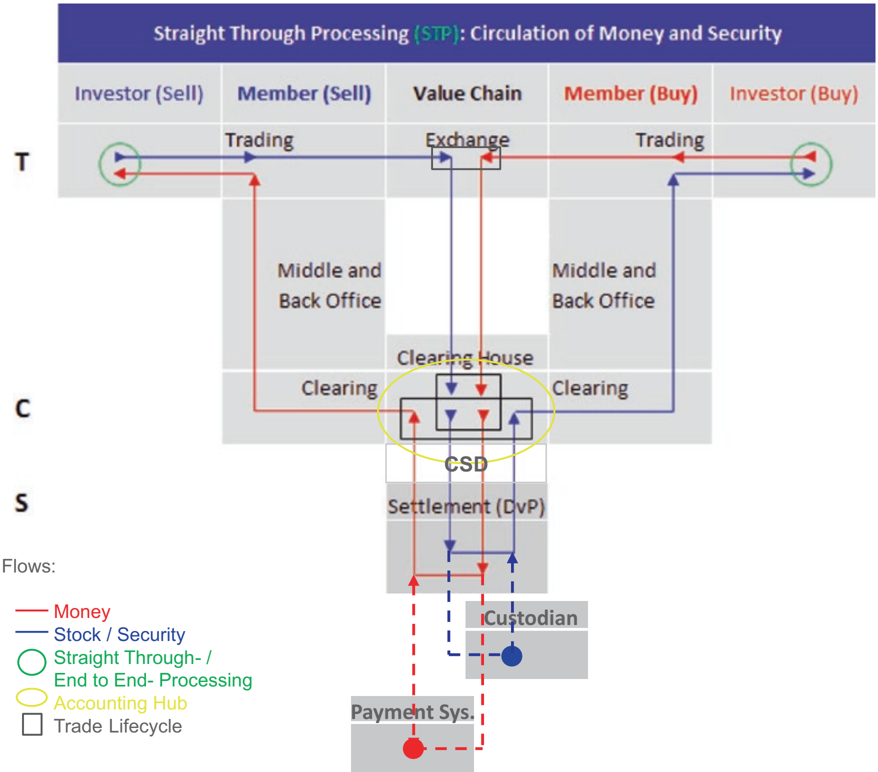


Fig. 2.31 Straight t-hrough processing of money and securities (schematic)

2.7 The Role of Technology Within the Value Chain

2.7.1 Key Characteristics

Today's equity markets are fully automated, and transactions are seamlessly processed along the value chain by customized, purpose-built IT systems. Each step in the value chain typically involves different systems: for example, market data dissemination, trading, clearing/risk, settlement, and custody. Organizations which cover more than one of these steps may deploy tightly integrated systems across their part of the service spectrum.

However, even when systems are operated within a single entity, the functional components are still fairly distinct and self-contained. Efficient transaction processing requires well-defined electronic interfaces in order to avoid manual intervention and to allow for an uninterrupted process flow (STP). When the value chain spans multiple organizations—e.g., a stock exchange, a clearing house, a CSD, and a custodian—the handover of transactions between their respective IT systems needs to follow clearly structured confirmation protocols and must pass through distinct demarcation points, control processes, and time stamping.

While these systems tend to be reasonably well interlinked nowadays, they still display widely varying characteristics. Some of these key characteristics apply to the specific business functionality, whereas others impact the underlying technology (Fig. 2.32).

Depending on the step in the value chain, a different set of key characteristics is relevant. For example, the functional richness of a risk management system far exceeds that of a trading engine. Or, while settlement data should most certainly be kept in a persistent environment, market data, however, can become fairly irrelevant as soon as it is superseded by new transactions and the market has moved on.

2.7.2 The Trading Platform: The Formula One of Exchange Systems

The IT systems used at each step along the value chain are typically based upon proven infrastructure components which are well established in the financial industry. While the functional richness and the complexity of the business requirements may vary, most components can be deployed on standard computing systems and use industry standard architectures.

This is not the case for those systems that cater to today's extremely fast response times in electronic trading. High-performance trading systems nowadays deploy technology that goes far beyond conventional designs to reduce transaction latencies to an absolute minimum. Leading-edge technologies, such as remote direct memory access, field-programmable gate arrays, and microwave transmission, can bring down order transmission and processing times down to microseconds (a millionth of a second). In Chap. 6, we outline specific design principles for these high-performance trading systems.

Strategic Importance of Key Information Technology Characteristics along the Value Chain

Business					Technology					
Functional Richness	Network Effect	Trans-parency	Fairness	Value Chain	Realtime Availability	Deadline Driven	Persis-tence	Low Latency	Computing Power	Capacity
–	–	✓	✓	Market Data Dissem-ination	✓	–	–	✓	–	✓
–	✓	✓	✓	Trading	✓	–	–	✓	–	✓
✓	✓	–	–	Clearing, Risk	✓	✓	✓	–	✓	✓
–	✓	–	–	Settlement	–	✓	✓	–	–	✓
✓	✓	–	–	Custody	–	✓	✓	–	–	✓

Fig. 2.32 Strategic importance of key IT characteristics along the value chain

2.7.3 Price Discovery: Software vs. Software

The price discovery process is defined by means of algorithms in the form of software. It is the objective of matching as part of the central system to determine the execution price,[24] sequentially for a specific situation as well as for continuous trading. The application of an abstract concept (like the principle of highest executable volume) via software and hardware secures the determination of one execution price per listed share at any time within the central order book. Software facilitates the sequential execution of a distinct, unmodifiable description of a process—matching—in numerous finite steps (Fig. 2.33).

The matching algorithm has the following properties:

I Syntax (form):

- Expressed in a specific language, formal and consistent
- Finite number of steps until a definite result is reached
- All steps are executable and sequential[25]
- Completeness, i.e., all possible "what if" cases are conclusively covered[26]

[24] Each specific situation in the CLOB has exactly one execution price.

[25] For price discovery only in technologically determined exceptional cases parallel.

[26] In philosophy an axiomatic system has to be consistent, whereas price discovery as a system has to be consistent AND complete (cover all possible cases).

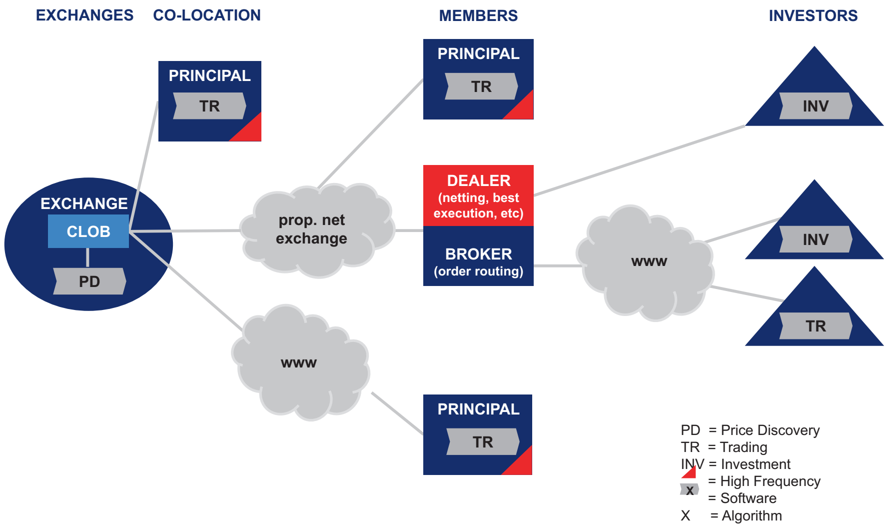


Fig. 2.33 Price discovery and order execution: software (PD) vs. software (TR/INV)

II Semantics (contents):

- Axioms or rather assumptions as unprovable elementary principles, describing an ontology.[27] Regarding price discovery, the most important principles are the principle of highest turnover with lowest surplus, and the principle of price-time priority.[28] For market participants the principle of equal treatment is of utmost importance.
- Consistency: axioms must not contradict themselves, while the same is true for the process.
- Completeness, i.e., all the axioms together contain all information to cover the entire price discovery process. Any further additions would lead to inconsistency.

2.8 Risk Management Along the Value Chain

One of the key functions of an exchange organization (along with trading, clearing, settlement, and custody) is the provision of efficient risk and securities management services to participants in the international capital markets around the globe. Therefore, it is especially important that the exchange organization should also have appropriate procedures in place to protect itself from risk.

An exchange organization could categorize risk into three types: operational, financial, and business risks. Operational risks include system availability risks,

[27] Cf. [7].

[28] Cf. Chap. 4.

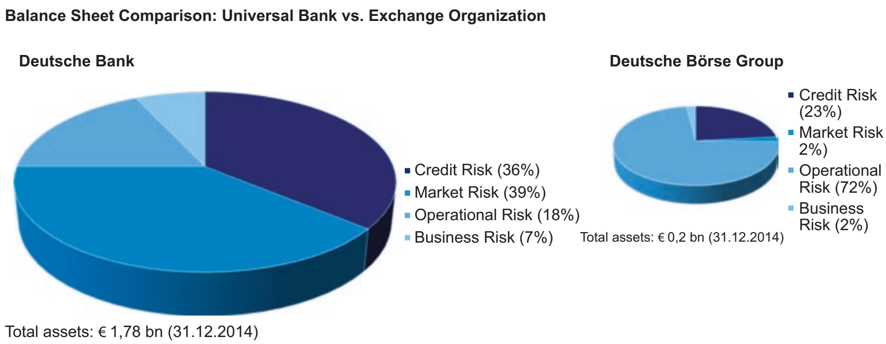


Fig. 2.34 Risk profile universal bank vs. exchange organization

legal and regulatory risks, and processing errors. Financial risks include primarily credit but also market and liquidity risk. Business risk reflects potential impacts on the organization's operating result attributable to economic, competitive, regulatory, or other market developments.

While the trading, clearing, settlement, and custody division of an exchange organization deal with the enormous volumes that are being transacted in the international capital markets, they avoid holding market positions that are at risk. Furthermore, credit risk is managed through strict customer acceptance and credit approval processes. Thus, while the exposures to customers can amount to billions of Euros, Dollars, Renminbi, or one of the other currencies in which the respective organization processes transactions, such exposures are generally to highly rated counterparties, and they are fully secured by high-quality collateral. In addition, the clearing house's exposure is protected by member contributions to the clearing fund while, in the settlement business, exposure is very short-term (primarily intraday). Only by protecting itself can an exchange organization protect its customers.

Thus an analysis of, e.g., Deutsche Börse Group's risk profile shows a significant difference from other financial institutions. While credit and market risk may account for 75 % of the economic capital requirements of a large universal bank,[29] at Deutsche Börse Group, credit risks account for only around 23 % and market risk for just 2 % (Fig. 2.34).[30]

2.9 Contractual Relationships Within the Value Chain

Describing the value chain of an equity market one must consider the trading layer, the clearing layer, and the settlement layer. These layers are integrated by the STP of transactions on a technical level. In this section of the chapter, we focus on the legal relationships that are required (or that at least are usually established) for each layer in order to execute and process securities transactions along the value chain (Fig. 2.35).

[29] Deutsche Bank quarterly report (30.09.2014).

[30] Deutsche Börse Group (31.12.2014).

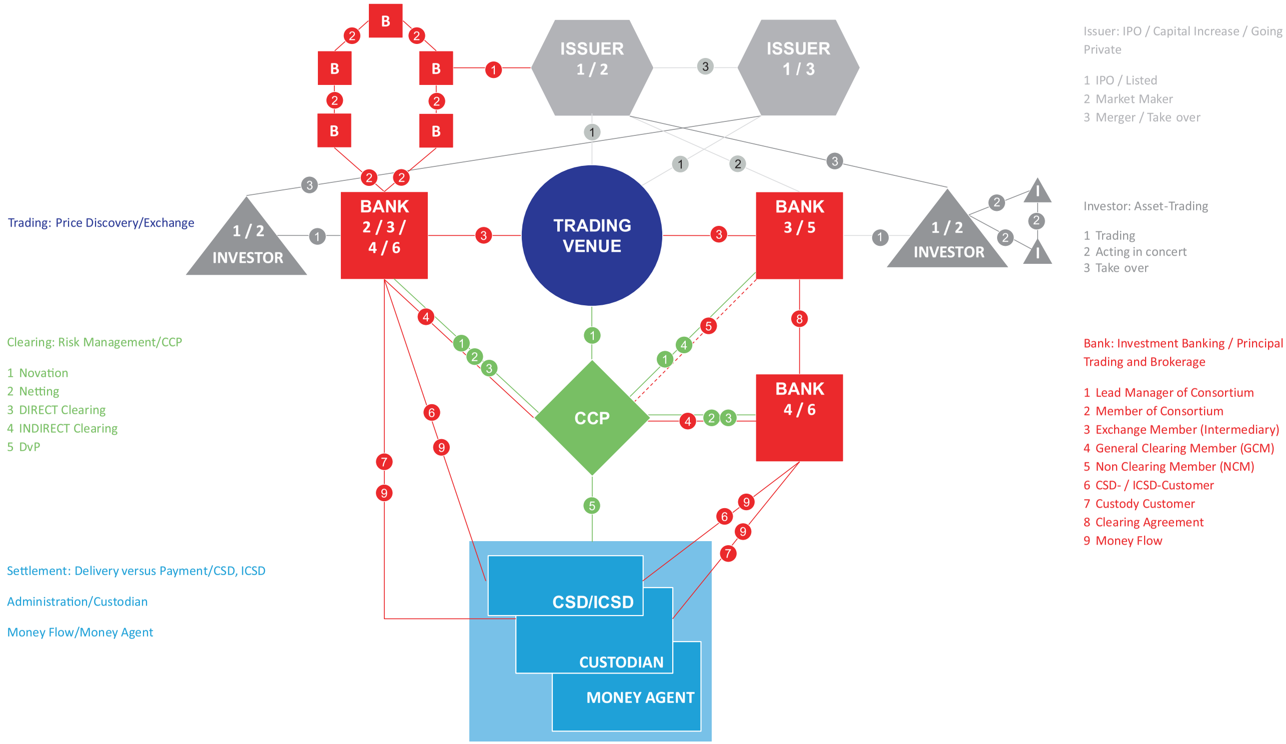


Fig. 2.35 Legal relationships within the value chain

- On the *trading layer*, i.e., on the level of the securities exchange as an organized marketplace, transactions are concluded based on an approved set of rules, and are supervised by the bodies of the exchange and the competent regulators. A precondition for the conclusion of transactions on the exchange is, on the one hand, the availability of tradable securities and, on the other hand, the admission of trading participants who have access to the trading system of the exchange, and who enter orders (or quotes) into the trading system for the purchase or sale of such securities. Therefore, on the trading layer, legal relationships are established by the exchange with issuers of securities admitted to trading on the exchange and with trading participants. As a rule, securities exchanges have no direct legal relationships to investors who are not at the same time issuers or trading participants.
- As previously discussed in this chapter, on most of the securities exchanges, a CCP is contractually interposed between the trading participants (*clearing layer*). The exchange, in cooperation with the CCP, may determine that transactions in specific securities are not eligible for clearing and that these transactions have to be settled bilaterally between the counterparties without the CCP being involved. Unless such bilateral settlement is foreseen, transactions in securities executed on the exchange are being cleared through the CCP who provides post-trade anonymity, netting, and counterparty risk management services. To clear transactions through the CCP, exchanges require that their trading participants have CCP arrangements in place to ensure the orderly settlement of transactions. They may choose to participate in the clearing of transactions as clearing members of the CCP or as non-clearing members, by facilitating third-party clearing members. On the clearing layer, the CCP establishes legal relationships to the entities that are involved in the clearing process, i.e., to both clearing members and non-clearing members.
- On the *settlement layer*, in order to settle the transactions executed on the exchange, CSDs organize the exchange of cash and securities on a delivery-versus-payment basis, and for that purpose they maintain technical interfaces with exchanges and their trading participants. For transactions that are cleared through a CCP, interfaces are in place with the CCP and its clearing members as well. For the settlement of transactions, legal relationships are established between the CSD and its customers. Such customers of the CSD are trading participants, intermediary banks, clearing members, and the CCP, depending on:
  - Whether transactions are settled bilaterally or are cleared through a CCP
  - Whether the settlement solution is implemented by the trading participants of the exchange

As clearing services by the CCP comprise the collateralization of transactions, the CCP is served by the CSD in particular, as the CCP needs legal and operational access to collateral locations.

In addition to the legal relationships addressed in the following sections, further legal arrangements are required to ensure that transactions are processed on an orderly basis for each layer, and along the entire value chain of the equity market.

In particular, on each layer agreements are concluded for the provision and/or operation of the required IT infrastructure. Such agreements are not necessary if the IT infrastructure is provided and operated by the securities exchange, the CCP or the CSD. Usually, contractual arrangements are also in place between the layers that provide for the use of clearing and settlement services and operational details of the processing of transactions across the different layers.

Whereas the trading, clearing, and settlement layers are integrated by STP on a technical level, from a legal perspective they can be integrated or linked together in various ways. Given that the exchange, the CCP, and the CSD may be independent, or that they may be operations within the same organization, two basic models exist. In the *vertical model*, the securities exchange, the CCP, and the CSD are all entities that are totally or substantially owned by the same company or group. This model is usually highly efficient and cost effective, but it gives only limited choice to the trading participants of the exchange with respect to the clearing and settlement of transactions. In contrast, the *horizontal model* separates the business into three layers: trading, clearing, and settlement. In this model, exchanges may integrate more than one CCP or CSD for the clearing and settlement of transactions. Though this model provides more choice for trading participants, from a legal and operational perspective it is more complex than the vertical model. Therefore, in the horizontal model, besides cost efficiency, it has to be taken into account that the implementation of different clearing or settlement solutions must not have an adverse effect on the orderly trading on the exchange, and on the orderly clearing and settlement of securities transactions.

2.10 Market Regulation: Investor and System Protection

Economic growth relies heavily on a safe and functioning financial system. The Financial Crisis revealed considerable shortcomings of financial market regulation and supervision; these were on the legislative side (namely regulatory and supervisory gaps) and on business side (ineffective risk management and a lack of transparency). This resulted in systemic weaknesses of the global financial system.

Based on these findings, the G20 Pittsburgh Summit in 2008 introduced a reformation of the regulatory environment of financial markets with the objective of establishing a "new normal" for the global system. While overriding principles for regulation and supervision remain the same, namely:

- Investor protection: including transparency, fairness, and equal treatment of all market participants
- System protection: focusing especially on operational and financial risk management

The focus of this new order was set to be safety, integrity, efficiency, and burden-sharing to prevent a recurrence of those events that initially triggered the crisis.

The Pittsburgh Summit also saw the introduction of the G20 regulatory reform agenda that defined the new scope of international rule-making and oversight.

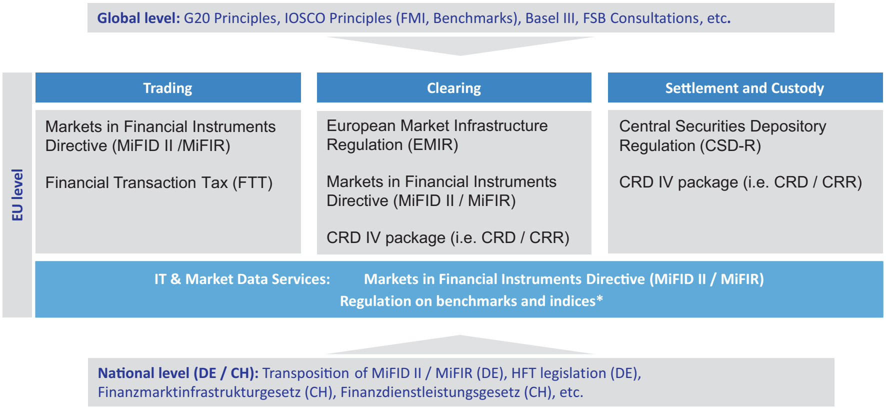


Fig. 2.36 European regulatory initiatives with impact on the value chain

Recently revised, the agenda currently comprises topics like how to end too big to fail, how to manage the shadow banking system, and how to increase the security of the derivatives market.

Due to the extensive interconnectedness of the financial system, the inevitable precondition to reach these goals is international cooperation in determining how best to develop a level playing field regarding supervision and regulation. Clearly, consistent implementation procedures are needed.

Especially for Europe, a joint federation of several states that had differing implementation practices pre-crisis, accomplishing all of this is an immense challenge. The same challenge exists for financial market infrastructure providers who are active in this environment. Most of the relevant initiatives covering European financial market regulation are developed on European level. Figure 2.36 outlines the number and diversity of regulatory initiatives that have had an impact on integrated exchange organizations like Deutsche Börse Group. All of them were developed following the principle of maximum harmonization (i.e., leaving little or no opportunity for member states to decide on specifications).

Although a very precise G20 process aims at a level playing field for regulation and supervision, a concrete tendency regarding regulatory arbitrage and fragmentation (especially between the EU and the USA) is becoming increasingly visible.

Different to the USA (where a majority of financial market regulation is covered by the Wall Street Reform and Consumer Protection Act, commonly known as Dodd-Frank Act), the European Union has developed numerous regulations covering different pillars of the financial system.

While Dodd Frank covers trading and clearing, EMIR, the European Market Infrastructure Regulation, regulated clearing only. Due to the prevailing variety of national regulatory systems, its realization takes much longer in the EU than it does in the USA, thereby allowing US financial firms to offer services faster in Europe than their European competitors. Another example is the prevailing intention of some European states to enact a financial transaction tax, which the USA completely rules out.

The coming years will see even more work on an international level to foster a consistent political strengthening of regulated markets and market infrastructure that make a substantial contribution to stable, transparent, and crisis-proof financial markets. Financial market infrastructure providers are not only facing challenges from regulation. Due to the overall transformation of the financial service industry, exchange organizations in particular can derive considerable opportunities from new customer needs and new business opportunities.

2.11 New Customer Needs

2.11.1 Scope and Mechanisms of the Financial Market Reform

The financial market reform is centered on preventing excessive risk taking on global capital markets by ensuring the transparency of derivatives marketplaces, by curbing leverage, and by imposing rigorous risk management. Consequently, financial reform targets market participants both sell side and buy side, as well as the venues and infrastructure through which they interact. Both elements are interrelated and shape customers' needs in the post-crisis environment. Although the dislocation originated in the mature Western markets, the reform agenda that is based on the G20 objectives applies globally, through either international accords like Basel III/BCBS IOSCO or concurrent regional frameworks like EMIR/Dodd-Frank.

The impact of regulation on customer needs is essentially threefold:

- First, regulators mandate the use of multilateral trading systems (MTS) or CCPs, and they selectively prohibit specific activities like prop trading.
- Second, within the range of licit market structures and roles, capital and liquidity rules influence the profitability of capital market business lines, and they provide incentives for the usage of standardized instruments and central clearing. As a result of these two factors, market participants align their business models and their balance sheet structures with the revised set of available transactional options and incentives. This in turn affects their customers' product choices and the distribution channels.
- Finally, market participants need to step up their capabilities in managing data, risk, collateral, and liquidity in the new regulatory environment. Buy-side customers are compelled to set up a front-to-back access infrastructure to cater to the multipolar market structure.

2.11.2 Mandated Change of Market Structure and Participants' Roles

The prescriptive elements of financial regulation define the range of permitted activities as well as eligible trading and post-trade platforms. The Volcker rule that practically eliminates prop trading by US banks has partly shifted these activities away from US banks and into hedge funds that access markets and financing through prime brokers rather than directly as sell-side actors. **High-frequency trading** in Europe will likely be restricted, but not eliminated, by upcoming EU regulation.

Moreover, regulators define and mandate the use of transparent, MTS for specific standardized instruments, thereby displacing models with dealer-centric liquidity provision. Listed exchanges, dealer-controlled entities, as well as network- and technology-focused players have competed for operating these nascent marketplaces. Central clearing has been imposed for a large range of derivatives contracts. These rules will be complemented by obligatory margining for non-centrally cleared derivatives.

Finally, a comprehensive registration obligation for OTC derivatives has been put in place, but this does not usher in a new way of transacting. Market participants need to build up connectivity to electronic platforms and to accommodate new reporting, trading, and clearing processes. The latter affects certain nonfinancial entities that are using derivatives vis-à-vis to direct market participants.

In summary, post-crisis regulation redefines market models and transaction processes, affecting direct and indirect market participants. Some former activities of the dealer community are displaced or shifted to less regulated entities.

2.11.3 Strategic Balance Sheet Management

Banks need to manage their balance sheets more strictly in light of Basel III's increased capital and liquidity standards. The revised capital requirements drive the profitability of capital market business lines and, consequently, the allocation of risk capital across asset classes, product types, and trading/post-trade channels. Channels are relevant as CCP positions enjoy lower capital requirements (lower counterparty weight and CVA exemption) than equivalent exposures to other counterparties where both routes are permitted. Product choice matters as—among those positions held at CCPs—easy-to-liquidate standardized derivatives are assigned lower margin requirements than typical OTC contracts.

The benefits of standardization and clearing enter the equation of product profitability. Regulation accentuates the trade-off between profit margins and capital costs. Since dealers are compelled to charge higher costs to their clients, they are incentivized to opt for plain vanilla products and to accept more basis risk in return for lower fees. Finally, the prospective Basel III leverage ratio will add a channel-agnostic dimension to capital costs, and the impact will depend on its parameterization. However, the design of this rule will first put low-margin trading businesses at risk.

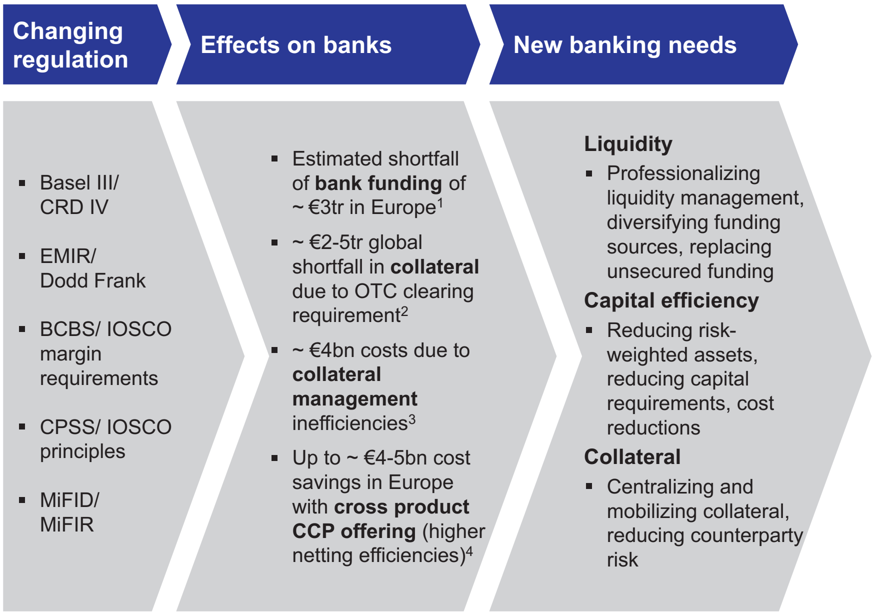
In light of the revised cost structure of the banking/broker-dealer sector, the provision of liquidity will be reorganized. For instance, bond dealers may cease to provide liquidity as the increasing capital cost of inventory erodes the profitability of market making in a low-yield, low-volatility environment. Investors will therefore have to seek liquidity in alternative multilateral market setups, or in the shadow banking sector that involves less regulated players like hedge funds. This brings new, technology-focused operators to the fore in fixed income, and it modifies trading logic and infrastructure.

Apart from capital requirements, Basel III comprises specific regulations on short- and mid-term liquidity, which calls for a proper structuring of the balance sheet, and a robust toolset for operational liquidity management.

In summary: The capital regulation of Basel III, which was targeted by regulators for the sake of systemic security, has crystallized the migration towards central trading of standardized instruments and clearing. In this process, Basel III reinforces the reallocation of liquidity pools away from dealers, and it further increases the scale and scope of central counterparty clearing.

2.11.4 New Requirements: Market Access, Liquidity, and Post-trade

The partial retreat of the dealer community compels the buy side to build up an infrastructure to access directly standardized products traded on multiple venues. They need capabilities to ensure best execution across fragmented liquidity pools, and to manage the risks inherent in the use of imperfect hedging tools that are provided by the wholesale markets of standardized derivatives. As these requirements add to clearing obligations, major buy-side players will have to match the dealer community's trading and post-trade technology (Fig. 2.37).



1) Quantitative impact study of Basel Committee on Banking Supervision (December 2010)
2) Celent study "Cracking the Trillion Dollar Collateral Optimization Question" (August 2012)
3) Accenture and Clearstream study "Collateral Management" (2011)
4) "The future of central clearing" study by Eurex Clearing and Oliver Wyman (April 2014)

Fig. 2.37 Changing regulation drives new customer needs

For the dealer community, the increasing centralization of counterparty and market risks within CCPs dramatically enhances the effectiveness of risk management. It facilitates timely and comprehensive transparency across the market, and it increases the speed and robustness of risk mitigation as well as recovery and resolution processes. At the same time, high-grade collateral is becoming an increasingly scarce resource as initial margin requirements—for CCP or bilateral positions—drive up costs by adding to the funding-related collateral requirements of central banks and other liquidity sources.

Market participants will best deal with this challenge if they are enabled to draw on a virtually integrated pool of their eligible assets, and can allocate these flexibly to risk exposures or liquidity providers. CSDs and global custodians provide collateral management services for that purpose. Both are complementary in principle, while their division of labor differs in the USA, Europe, and other financial centers due to the specific market infrastructure arrangements.

In Europe, CSDs play a pivotal role in the collateral management and securities financing that they offer as complementary services to their core notary, safekeeping, and settlement functions. Central bank money access, and close collaboration with the CCPs, has enabled the design of integrated, high-security collateral management solutions that span trading and financing markets.

In contrast, collateral management infrastructure in the USA is centered on custodian banks, specifically J.P. Morgan and Bank of New York Mellon, which are both in the securities financing and derivatives clearing markets. Although the system has proved workable, assigning central infrastructure functions for cash and securities settlement to commercial banks exposes US markets to heightened systemic risks.

The post-trade infrastructure in financial centers other than Europe and the USA is currently undergoing upgrades in terms of rulebooks, processes, and tools. The process involves the replication of specific collateral management capabilities developed by infrastructure providers, often in cooperation projects.

2.11.5 Outlook: Transformation of Global Capital Markets

While the financial market reform applies globally, the magnitude of the transformation varies across regions. The US and European approaches have, up to financial crisis, built up oversized banking sectors and excessive leverage. Consequently, the transformation of their capital markets in terms of downsizing and structural adjustments is far more pronounced than is the case in Asia. The speed of regulatory action is equally varied. The USA has implemented its "Wall Street Reform" in the Dodd-Frank Act more swiftly than the EU equivalents MiFID II/EMIR, as the decision-making process in the EU is invariably more complex. National regulators in Asia follow suit, but their frameworks are less pervasive in OTC derivatives rules in light of the limited size of these markets in their region.

This transformation has spawned a plethora of new reporting, trading, and clearing offerings to facilitate compliance and optimal operational management within the new landscape. For exchange organizations (reasonable) regulation represents:

- An asset due to positive effects on market integrity
- Opportunities for new products and services

Market infrastructure providers and other players from the financial market ecosystem have competed to serve emerging customer needs. The competition is global, but mutual access rules are still being fleshed out.

US operators have had a head start as they were in a position to upscale their businesses ahead of EU competitors. However, integrated European exchange groups are uniquely positioned to deliver innovative post-trade technologies by integrating cash, collateral, and risk management functions. Such evolution towards holistic post-trade solutions started before the crisis. This development was increasingly driven by the cooperative efforts of different market infrastructure providers and their respective regulators and central banks. The same is true for Asian markets as well. But, whereas in the USA and in Europe the regulatory structures are set on implementation, in Asia some key market players, primarily China, can still make strategic use of the second mover advantage.

Acknowledgments The author would like to acknowledge the valuable contribution of Mr. Ulrich Strohmeier and thank him for his continued support regarding Section 2.11 New Customer Needs.

References

1. U.S. Code Chapter 2B - SECURITIES EXCHANGES: http://www.law.cornell.edu/uscode/text/15/chapter-2B
2. Schwartz, Robert A. / Francioni, Reto (2004): Equity Markets in Action. The Fundamentals of Liquidity, Market Structure & Trading. Hoboken: Wiley.
3. John Losee, A Historical Introduction to the Philosophy of Science (Oxford; Oxford University Press, 2001) 228; cited according to Thomas S. Kuhn, The Essential Tension (Chicago; University of Chicago Press, 1977), 321–2.
4. Eurex Clearing / Deutsche Börse Group (2014): How central counterparties strengthen the safety and integrity of financial markets. Frankfurt / Main: Deutsche Börse AG.
5. Clearing Conditions for Eurex Clearing AG: https://www.eurexclearing.com/blob/clearing-en/51612-136778/238376/34/data/clearing_conditions_en_ab_18_11_2013.pdf.pdf
6. Clearstream / Accenture (2014): Collateral Management. Unlocking the Potential in Collateral: http://deutsche-boerse.com/INTERNET/MR/mr_presse.nsf/0/DE08E32FA7CC4EFBC125790B004A79F4/$File/CM_PoV_Clearstream_Collateral_Management_sept2011.pdf?OpenElement
7. Gómez-Pérez, A./ Fernández-Lopez, M./ Corcho, O. (2004): Ontological Engineering. Berlin: Springer
8. Deutsches Börsengesetz: https://www.gesetze-im-internet.de/bundesrecht/b_rsg_2007/gesamt.pdf
9. Börsenordnung für die Frankfurter Wertpapierbörse: http://www.deutsche-boerse-cash-market.com/blob/1198492/a29dc9dfa533193a4ef54319da7f63ae/data/2016-03-18-Boersenordnung-fuer-die-frankfurter-wertpapierboerse.pdf

Chapter 3
Primary Market: Bringing Products to the Market

Cord Gebhardt and Jan Strecker

Savings and investments between suppliers and users of capital are channeled by the capital markets from retail and institutional investors to businesses, government, individuals, and others. Capital markets are vital to the functioning of an economy since capital is a critical component for generating economic output. Capital markets include the primary markets that sell new stock and bond issues to investors as well as the secondary markets for the trading of existing securities.

3.1 An Initial Public Offering

An initial public offering (IPO) is a type of public offering of stock in a company in which shares are sold to investors and a private company is transformed through this process into a public company. IPOs enable companies to raise capital for expansion, to potentially monetize investments of early private investors and to become publicly traded enterprises. After the IPO, when shares trade freely in the open market, money is then exchanged for the shares between public investors in the secondary market.

3.1.1 Issuing Business

The objective of an issue is the creation of new fungible securities, legally and freely transferable and tradable. In practice, the issuing business is conducted predominantly by credit and financial service institutions, that is, by issuing banks.

C. Gebhardt (✉) • J. Strecker
Deutsche Börse Group, Mergenthalerallee 61, Eschborn 65760, Germany
e-mail: cord.gebhardt@deutsche-boerse.com; jan.strecker@deutsche-boerse.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_3

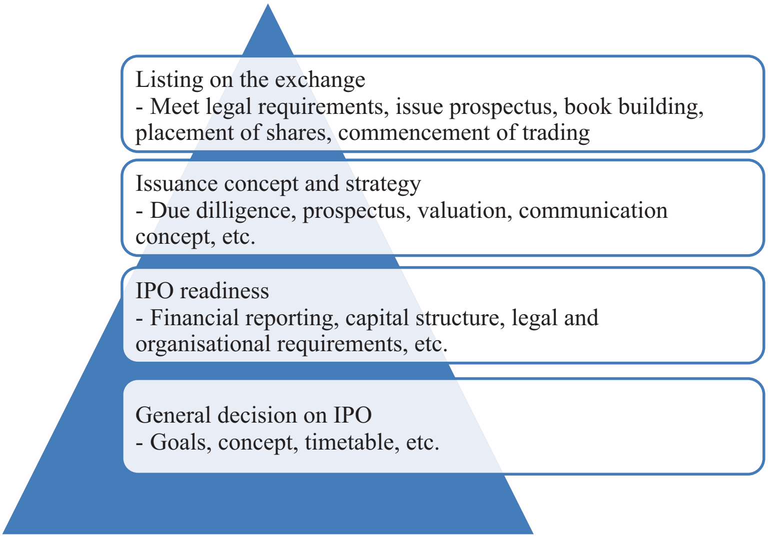


Fig. 3.1 IPO process

This covers their participation in the security-issuing process, especially in the advising, preparation, and placement of the issued securities.

There are numerous legal regulations for both the issuer and its advisors within the issuing process (Fig. 3.1). An issue of securities, namely the issuing business, does not necessarily involve a stock market listing of the newly issued securities. The issuance is, in general, completed with the placement of the new securities, i.e., their selling to the investors in the primary market. Nevertheless, a stock exchange quotation, resulting in the tradability of the newly issued and placed securities on a stock exchange in the secondary market, may benefit the issuer. As a rule, investors have an increased interest in a stock exchange quotation based on the simple proposition of exchange trading possibilities. These include price transparency, liquidity, supervision, and legal certainty of trading as well as the "information duties," which must be continuously fulfilled by the issuer. For example, there is financial reporting and ad hoc disclosure. Furthermore, capital investment and insurance companies as "institutional investors" are subject to numerous legal acquisition restrictions in the case of non-exchange quoted securities. Therefore, in practice, a share issue frequently takes place along with a stock market listing.

3.1.2 Placement of Shares

At a public offering, the new shares are offered to an unknown variety of investors for purchase. Within the scope of a public offering, different methods of placements have to be distinguished. In this respect, the tender or book building procedure is of major importance. Here, investors are requested to launch a generally discretionary bid for the new shares. Therefore, no fixed price or offering price is targeted by the issuer and the issuing bank. Instead, the price is determined by supply and demand in the capital markets, though in practice a minimum price is at least often specified.

Besides the tender or book-building procedure, the “public subscription” and the “market sale” are common. Generally, the public subscription is different from the tender or book-building procedure because, unlike the former, an offering price for the new shares is fixed. Consequently, investors may only launch the purchase bid at a predetermined price. The market sale is characterized by a placement of the new shares allocated to the members of the issuing banks with their respective customers.

At the end of a placement via a public offering, the “allotment” or “allocation” follows with the distribution of the new shares to the bidding investors. Furthermore, the allocation has central legal importance: The individual purchase agreements are concluded only with the allotment of a specific number of shares to each purchasing investor. Additionally, the allocation is of special importance if the demand for new shares exceeds the offering, meaning the investors have launched more purchase bids than new shares are available in the oversubscribed issue. In such cases, not all investors may receive their individually requested amount of shares. Generally, a partial or quota allotment occurs in a process known as “scaling down.”

Both the issuer and the issuing banks have wide discretion in the determination of concrete allotment criteria. Investors are not entitled to claim equal treatment in reality; however, the allotment is often performed in an arbitrary manner based on the business interests of both the issuer and the issuing banks. From an issuer point of view, this includes the composition of the shareholder base while the issuing bank is usually taking client relations into account.

The price investors pay is of paramount economic importance, especially for the issuer in a new shares placement. In practice, the “open pricing” and the “fixed pricing” face each other. In the first case, there is no firm price—at most, an indicative or floor price that is determined by the issuing banks; in the second, a placement is performed at a determined price from the start.

Finally, from a practitioner’s viewpoint, the high degree of internationalization of the issuing business must be highlighted, especially the dominance of Anglo-American conventions in all areas. The growing interconnectedness of banks is not the only reason though. Another is that increasingly issues are placed internationally—in a number of countries—for both issuers registered abroad offering their shares domestically and, conversely, for domestic issuers actually placing their shares abroad. The evidence is clear. The language of prospectuses must comply with their controlling legal framework. English is the chosen language, not only in more prospectuses today but also in the international capital markets (the common language within the issuing business is heavily Anglo-American oriented). Finally, the international practice has mainly influenced the issuing business, for example the implementation of the book-building procedure within a share issue.

3.1.3 On- and Off-Exchange Trading

In principle, the issuing process is completed with the placement or the disposal of the shares to investors in the primary market. However, the shares may not be traded at an exchange in the secondary market. Generally, a separate procedure must be undertaken for admission to exchange trading. The exchange represents the legally regulated and, therefore, the public and organized part of the capital market. At the exchange itself, capital is not raised, but securities are traded.

The public offering and the admission of the shares to exchange trading must first be distinguished. A public offering is legally defined and generally covers each and every message to the public that contains sufficient information on the offering conditions and the offered securities. This enables an investor to decide on a purchase or subscription of such securities. A public offering and admission to exchange trading may be combined. This means that shares are publicly offered as well as admitted to exchange trading. If the issuer offers its shares to the capital market for the first time, it is called "going public," an "initial public offering" or "IPO."

In Europe, the rules for a public offering as well as the rules for admission of securities to exchange trading have been harmonized by EU legislation. The goal of this far-reaching step is the creation of a single market for securities, especially for investor protection and market efficiency. On the one hand, EU-wide harmonization of prospectus information ensures equal investor protection and on the other a consistent requirement profile or a level playing field for issuers. Most important, the introduction of the "European passport," the reciprocal recognition of rules and regulations, implies that issuers may execute cross-border issues with only one prospectus at the same time.

The most important part of the German implementation of this EU regulation is the Securities Prospectus Act, which contains the following key aspects: The previously separated competence for prospectus examination between the exchanges of the German federal states and the federal authority, the Federal Financial Supervisory Authority (BaFin), has been unified. Central authority for any prospectus examination is now the BaFin. Therefore, the exchanges in Germany are generally no longer competent to assess a prospectus. The approval of a prospectus permits both the public offering of the respective securities and the application for trading at an organized market, i.e., exchange trading. Nevertheless, the formal distinction between the prospectus approval and the admission to trading is maintained. Therefore, approval of the prospectus does not automatically permit admission to exchange trading. The exchanges continue to be the competent bodies in the decision-making process for the admission to trading—and a consequent listing—independently of the BaFin.

The creation of a "European passport" typically leads to EU-wide validity of an approved prospectus. That is because a prospectus approved within the country of origin of the issuer is legally valid in all other member states and does not require a new examination by the other national authorities.

The EU-wide content-related harmonization of the prospectus requirements has mandatory prospectus formats. For instance, the deadline for prospectus approval is harmonized across the EU and, therefore, grants broad equality of procedure. Accordingly, as a rule, BaFin must inform its decision within 10 business days or within 20 days in the case of an IPO.

It must be noted that the EU-legal harmonization of the prospectus system does not affect the current prospectus liability. Both the question of prospectus responsibility (who is liable) and the question of the scope of liability (being liable for what and to what amount) still have to be answered according to the respective national laws of the single member states. An EU-wide harmonization of the different prospectus liability regimes has not been achieved yet.

In Germany, the exchange trading of securities at the regulated market requires an administrative admission (Listing). In legal terms, the admission to exchange trading is a decision governed by public law. Furthermore, admission to trading has to be separated from the quotation at the later introduction of the admitted securities. Legally, quotation is defined as the first price fixing at the exchange.

However, the exchange trading of securities at the unofficial regulated market (Freiverkehr) at the exchanges does not require an administrative admission. Trading here requires only an “inclusion” to the unofficial regulated market on a private law basis. Legally, it is a contract between the exchange operator and a trading member, typically without any participation by the securities’ issuer. The regulated market of the exchange is an organized market regulated and supervised by state bodies. The unofficial regulated market (Freiverkehr) of the exchanges is part of the off-exchange trading (over-the-counter (OTC) market). This is mainly organized under private law and subject to only limited state supervision.

In Germany, there are currently eight stock exchanges, each legally equivalent and independent public law institutions: two in Berlin, and one each in Düsseldorf, Frankfurt, Hamburg, Hannover, Munich, and Stuttgart. If shares are supposed to be traded at more than one German exchange, then separate admission procedures have to be executed. These are usually subject to the same legal requirements. Formally, filings are separately completed per exchange. State supervision of the German exchanges is decentralized by the applicable German federal state where an exchange must have a registered seat. BaFin as central state authority for financial services generally has no supervisory functions for the exchanges.

3.1.4 Advantages and Disadvantages of a Listing

Commonly, it is at the issuer’s full discretion to decide if the issued shares are supposed to be traded at an exchange. In practice, this decision is a multilevel process with the contribution of numerous external advisors, such as banks, auditors, lawyers, investor relations advisors, and marketing agencies. This usually starts several months before the formal listing procedure begins. At the heart of the decision is weighing the advantages and disadvantages connected to a company listing.

A listing offers the following advantages:

The issuer's publicity and level of public profile increase. Consequently, this may lead to an increased demand for the issuer's products and services as well as an increasing attractiveness for qualified employees. Moreover, the issuer's reputation may be boosted, especially with the inclusion in an exchange index. Investors are now able to trade the purchased shares at a legally regulated and supervised market at any time (for the most part), and at transparent pricing due to public trading. Furthermore, in practice, options for issuing additional capital market instruments (e.g., bonds) have to be named along with the creation of an "acquisition currency" (one's own listed shares could play this role) for investments in payment in asset deals.

Besides these advantages of a listing, being in the "spotlight" may create disadvantages as well. Significantly, this includes the risk of a takeover by a competitor buying up the issuer's exchange-traded shares. In addition, there is also a risk of granting extensive participation rights to external third parties. That may even lead to a loss of entrepreneurial freedom in decision making by the "old" management; for example, there is the risk of dismissal and replacement of the current management by a supervisory board increasingly occupied by "outside" persons. It must also be mentioned that there is a certain analyzability of the issuer for competitors (transparent issuer) based on the legal publication requirements of a prospectus, ad hoc disclosure, and financial statements. Additionally, there are one-off and ongoing costs of a listing, e.g., publication requirements, annual shareholders' meeting, and analysts' meetings. Finally, and this should not be underestimated, the exchange presence causes a continued pressure on the issuer towards its investors to justify its existence. The issuer may be immediately "punished" for a failure to meet publicly announced goals, which would be reflected in a decline of its stock price. This, in turn, might influence the issuer's financial standing. There is also a certain danger that the issuer's management may adjust corporate decisions more for the short-term effect on the share price rather than the company's long-term benefit.

3.1.5 Delisting/Going Private

A company could, for various reasons, also decide to delist its shares from the market and become a private company again. The reasons for a delisting could be either involuntary or voluntary. An involuntary delisting could be the result of violating regulations and/or failing to meet financial specifications set out by the stock exchange. A voluntary delisting might be the result of different considerations altogether. A listing costs money to maintain, which may not be justified, particularly for smaller companies. In addition, recurring expenditure for financial reporting requirements, ad hoc disclosures, and investor relations and the increased demands on management to develop high-quality relationships with analysts and investors need to be considered. A delisting also frees the company from certain transparency and disclosure obligations. Finally, a delisting provides greater strategic and financial freedom for a company. For instance, a delisting will facilitate long-term strategic planning as short-term considerations become less important.

Restructurings may be conducted with less public attention. In the event of a takeover by a strategic investor, a planned consolidation or reorganization can be more easily effected.

3.2 Financial Market Communication

3.2.1 Equity Story

The equity story presents the company's core competencies, success factors, and future prospects.

The success of an initial public offering largely depends on the communication phase ahead of the listing. The central element of the capital market communication is the company's equity story. It includes all the main characteristics including the business model and the strategy, and it positions the company in front of potential investors. Thus, the equity story is a way of "translating" the company's strategy into the language of investors and analysts. Because of its importance for the IPO's success, the equity story requires detailed work, well ahead of the listing date. It involves the company's top management and external advisers.

An explanation of the company's business model in as simple and convincing terms as possible is an important element. By the same token, adapting an outside-in view on the company to reflect the view of analysts and investors is equally important because capital market participants usually take a peer group approach in evaluating companies. In this peer group approach, investors are comparing companies with the same business model for the most attractive investment opportunities in the sectors.

Ultimately, the equity story should answer the questions what the company is doing and why an investor should participate in the IPO. The development of the equity story can be conducted in three steps:

1. Analysis of sector trends, growth potentials, and peer group positioning
2. Analysis of the strength, weaknesses, opportunities, and risks of the company (SWOT analysis)
3. Preparation of the investment case and demonstration of the concrete benefits for the company of the initial public offering

The following elements are vital in any equity story:

3.2.1.1 Company-Specific Factors

1. The company's product and/or service offering (e.g., key factors for revenue generation, development costs, profitability, product life cycles, dependencies on suppliers or raw materials, seasonal influences)
2. Prospects and sustainability of the business model (e.g., strategic goals, growth prospects, use of IPO proceeds, external growth opportunities, potential shareholder returns)

3. The company's client base (e.g., size and structure of the client base, client behavior, dependency on large clients)
4. The company history and its management (e.g., market position of the company, financial track record, patents or intellectual property rights, experience of the management, access to qualified staff, M&A track record)

3.2.1.2 Sector-Specific Factors

1. Attractiveness of the sector (e.g., description of the sector and the competitive advantages of the company, sector studies including growth assumptions, consolidation scenarios)
2. Competitive position in the sector (e.g., peer group description, benchmarking of product and/or service offering, market shares)
3. Barriers to entry (e.g., unique selling propositions, technological advantages)
4. Regulatory environment (e.g., relevant laws and regulations, subsidies)

3.2.1.3 Other Factors

1. Corporate structure of the company (e.g., legal structure, shareholdings, related parties, corporate law history)
2. Existing shareholders (e.g., anchor shareholders, blocking minority)
3. Transaction structure of the IPO (e.g., remaining shareholders, amount of capital to be raised)

The IPO will only be successful if the company can deliver and communicate an attractive and sustainable equity story to capital market participants. It can be helpful if financial communications advisors and the bank consortium support the equity story development based on the company's experience. The operational know-how of the company can be complemented with this capital market experience. If the core elements of the equity story are developed, different communication formats need to address the specific requirements of the target groups, from institutional and retail investors to sell-side analysts and the media and other stakeholders. The equity story will be the core of the entire financial communication during and after the initial public offering process. The impact of a well-crafted equity story should not be underestimated.

3.2.2 IPO Communication

In the marketing phase prior to the IPO, the equity story is presented to potential investors.

The IPO communication process can be split into four parts: preparation phase, image-building phase, pre-offer phase, and offer phase (Fig. 3.2).

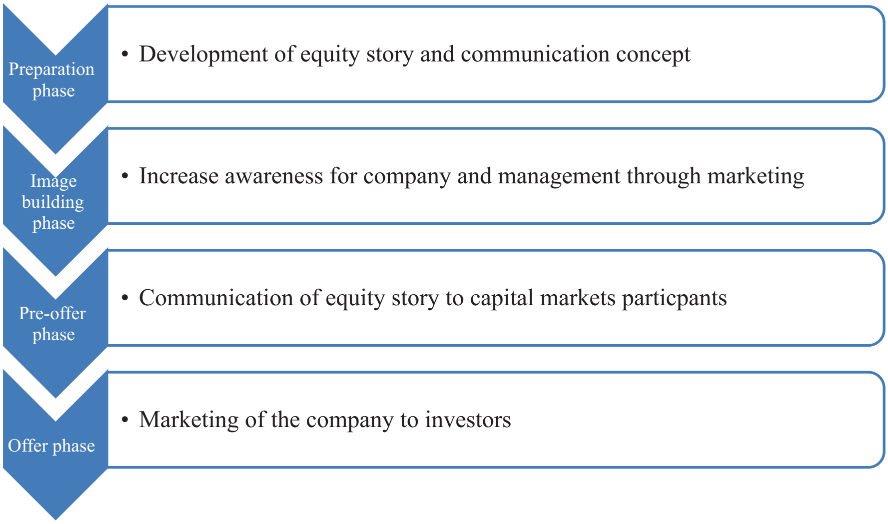


Fig. 3.2 IPO communication steps

In the preparation phase, the detailed communication concept for the IPO has to be developed. The heart of the concept is the equity story, serving as the basis for marketing the company to sell-side analysts and investors. As part of the communication concept, a clear plan is required with a view of the most important company stakeholders, of who should buy shares in the IPO process, and of who can serve as opinion leaders for investors and the wider public.

The image-building phase focuses on marketing the company and its products. In this phase the potential IPO is not actively mentioned, but the phase helps to increase awareness for the company and its management in the market generally, and with the media and investors in particular. Thus, this phase is especially important for companies less known to the public, or with more complex business models. The main marketing instruments in this phase are media relations work; company publications (e.g., annual reports); and image campaigns in newspapers. During the image-building phase, the equity story should also be presented to, and discussed with, sell-side analysts of the bank consortium. This is a further opportunity to collect feedback from capital market participants with a broad overview of the peer group. This also prepares the analysts to write the research reports for the IPO.

The pre-offer phase starts with the announcement that the company is preparing an IPO. At this stage, the company normally does not disclose timing details of the going-public date. The pre-offer phase allows the issuer to finally evaluate the IPO readiness and make adjustments to the equity story. In the pre-offer phase, all of the communication formats including financial reporting, investor and media relations, a website, and company presentations should be set up. In the pre-offer phase, the company can also ensure that there is sufficient sell-side research coverage after the IPO. This can be achieved by meeting with banks and research houses that are already covering the sector or peer group companies.

Research from sell-side analysts, who are opinion leaders, is an important conduit to increase penetration of the equity story in the capital markets. The media is similarly important. This applies especially to companies that are targeting retail clients with their equity story. Institutional investors can also be reached through news coverage.

The offer phase is kicked off with a conference for analysts and investors as well as media. At this point, the securities prospectus is publicly available and all details on the issuance structure are known. In an ideal world, the documents do not include major negative surprises for the market, such as an unexpected use of IPO proceeds and unknown risk factors, since this would undermine the company's credibility and its management. Following the conference, the company's management will meet investors during a road show.

Ultimately, institutional investors, including pension funds, hedge funds, and wealth managers, are the most important stakeholder group in the pre-IPO marketing process. They expect that company management presents the equity story ahead of the IPO in one-on-one or small group meetings. Those meetings serve as a basis for the investors to evaluate management's credibility. These investors expect a high degree of capital market orientation that includes transparent, up-to-date, and key company information. Moreover, institutional investors will be focusing on the quality and accuracy of the company guidance within the applicable legal framework. For investor meetings, a clear and crisp presentation should serve as the documentation. In addition, the company should try to anticipate as many questions as possible ahead of the meetings, drafting possible responses in a Q&A document.

The day of the listing is the highlight of the IPO process. This is also when all the measures for the unending investor relations activities must be ready.

3.2.3 Ongoing Investor Relations Activities

Key elements of an effective communication with investors are transparency, consistency, and credibility.

A successful IPO is an important milestone in a company's development. From the moment of the IPO, the company has a new product: its own shares. The shares need to be supported and promoted, like other products or services the company is offering. In this sense, the IPO is a landmark, the beginning of a new company era. The primary goals of investor relations are a fair capital market valuation, and an optimized cost of capital. The investor relations function can achieve these goals through a stakeholder-oriented, transparent, consistent, and credible financial market communication (Fig. 3.3).

Transparency: Early in the process, the company has to identify the key business and financial metrics it plans to provide to investors after the IPO. Special consideration should be paid to the metrics that peers use to describe their businesses and to provide guidance on their future performance. Understanding and adapting these standards will keep the company aligned with what investors are accustomed to receiving from the peer group of the issuer, and the issuer must demonstrate a commitment to transparent communications.

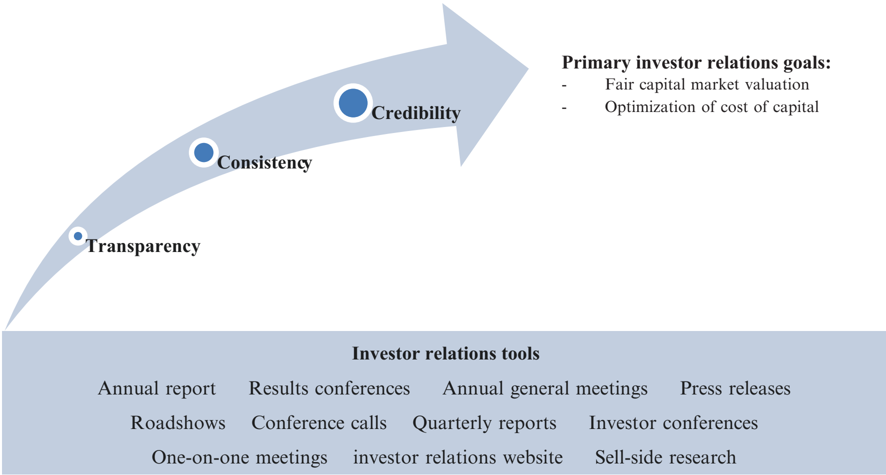


Fig. 3.3 Investor relations tools

Consistency: Investors are looking for new information in every interaction with the company. Any difference in messaging, content, tone, or frequency/timing of communications can be perceived as an indication of changes in the business or outlook that could affect the company's share price. Therefore, consistency in communications is very important. But the company must also be flexible enough to adjust to different business conditions.

Credibility: A major communications goals is building trust and credibility with capital markets' participants. To this end, credible financial market communication is essential to maintain the trust of investors in a long-term investment. The company has to inform capital market participants in an honest and fair way about the strategy, its objectives, and its business development. Furthermore, changes of relevant assumptions and parameters in the business model must be communicated on a proactive and timely basis. Greater trust among investors in a company normally results in higher demand for its shares; this might result in a valuation premium when compared to the company's peer group.

3.2.4 Investor Relations Tools

A wide range of mandatory and voluntary investor relations tools support a company to achieve the primary investor relations objectives.

Internet: The investor relations website is one of the first places for investors seeking more company information. Therefore, it should be user friendly, interactive, and easily accessible. Investors and analysts visiting the website must be supplied with all the information they need to conduct initial due diligence on the company, and to help them advance their investment decisions.

Results Releases: Reporting financial results to investors is an important medium for commentary about the business to the financial community. Typically, the earnings process includes a press release with financial data, an annual or quarterly report, and a conference call for sell-side analysts and institutional investors. These important communication formats help the company to demonstrate transparency regarding the way management speaks about the company's successes, challenges, strategy, and forward-looking guidance. Effective preparation is critical to management anticipating investor questions and proactively responding.

Investor Meetings: There are different formats with which management can engage with investors: road shows, where the company meets with institutional investors in one-on-one or group meetings; investor conferences, often with a presentation, followed by one-on-one or small group meetings for more detailed discussions; and company events such as analyst or investor days at the company headquarters, which allows access to the broader leadership team and company facilities. Webcasts are also increasingly used at investor events to increase global participation. Regardless of the format, these meetings provide valuable opportunities for context on financial results, explanation on strategy, and development of relationships with investors.

Media Releases: The company's media releases should be drafted with an eye towards what the content means for the business and how it will be perceived by the investment community. If applicable, media releases should tie news events to the company's stated strategy and show progress towards its objectives. If the announcement impacts the company's guidance for the quarter or year, these issues should also be addressed in the announcement. If the news has a high level of importance and complexity, a conference call for market participants can allow management to provide additional information on the event that prompted the media release.

The investor relation tools help the company to actively manage its shareholder base. Investors that are less well informed about the company, or whose investment style does not fit with the company, are more inclined to sell their shares. Diligent work is continuously required to identify the most important shareholders, to monitor changes in the investor base, and to engage investors in a dialogue to help keep them informed about the company.

Beyond this, there is also a continuous need to identify and attract new investors, the ideal target group being long-term-oriented investors. There will certainly be interest from sell-side analysts to help to market the company to prospective investors. But given the potential conflict of interests the sell side has, the company should be in charge of managing the investor base and targeting potential new investors.

Nevertheless, sell-side research analysts can help to increase visibility among investors. The company should therefore develop relationships with these analysts who play an important role in communications between management and the investment community. Still, prioritizing management's time with sell-side analysts can be challenging. Several factors should be considered by management when deciding which research firm to support: the quality of research, marketing events, and general opinion on the company. Finally, it is important to treat all analysts equally when interacting with the sell side in order not to give an information advantage to anyone.

Further Reading

Baruch Lev, Winning Investors Over, Harvard Business Review Press, 2012
Bosch/Groß, Das Emissionsgeschäft, bank-verlag, 2004 (German)
David Westenberg, Initial Public Offerings: A Practical Guide to Going Public, Practising Law Institute, 2014
Frederick Lipman, International and US IPO Planning: A Business Strategy Guide, Wiley, 2008
Jason Draho, The IPO Decision: Why and How Companies Go Public, Edward Elgar Publishing, 2006
Mario Levis/Silvio Vismara, Handbook of Research on IPOs, Edward Elgar Publishing Limited UK, 2013
Philippe Espinasse, IPO: A Global Guide, Hong Kong University Press, 2014
Steven Bragg, The Investor Relations Guidebook, 2nd edition, AccountingTools, 2014
Veit Etzold/Thomas Ramge, Equity Storytelling, Springer Gabler, 2014 (German)

Chapter 4
Secondary Market: Trading, Price Discovery, and Order Matching

Reto Francioni, Martin Reck, and Robert A. Schwartz

4.1 Price Discovery

4.1.1 Importance of Price Discovery

Price discovery is achieved as orders are submitted to a market and turned into trades. A transaction price is, of course, determined each time a trade is consummated, but *price discovery* refers to something more fundamental. Price discovery refers to the search for a value that best reflects the broad market's desire to hold shares of a stock. In economic parlance, price discovery involves the search for an *equilibrium value*. While price determination occurs on a trade-by-trade basis, price discovery is achieved only as a substantial set of orders is brought together, generally over a succession of trades.

An exchange's ability to deliver good price discovery depends on its market structure, namely the rules, procedures, and technology that define the exchange's trading platform (we delve into market structure later in this chapter). More efficient market structure enables the delivery of more robust price discovery for the broad market. The challenge, however, is formidable; equilibrium values are unobservable, they are continually subject to change, and they are not easily attained.

R. Francioni (✉)
University of Basel/Deutsche Börse Group, Basel, Switzerland
e-mail: reto.francioni@unibas.ch

M. Reck
Deutsche Börse AG, Mergenthalerallee 61, Eschborn 65760, Germany
e-mail: martin.reck@deutsche-boerse.com

R.A. Schwartz
Zicklin School of Business, Baruch College, CUNY,
55 Lexington at 24th Street, 268, New York, NY 10010, USA
e-mail: Robert.Schwartz@baruch.cuny.edu

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_4

An exchange's economic function is not simply to make the transactions but (and this is the bigger challenge) an exchange also "*produces*" *the prices* at which the transactions are made. Delivering quality price discovery is a defining function of a stock exchange. From a definitional perspective, any trading facility that has as its primary function the delivery of good price discovery can, de facto at least, be considered an exchange. Unfortunately, however, the price discovery function of an exchange typically receives insufficient attention in market structure discussions. This is largely attributable to the non-observability of equilibrium prices and, therefore, to the difficulty of quantifying the deviations of transaction prices from their equilibrium values.

4.1.1.1 Expectations

Market participants commonly believe that shares have *fundamental values*. The concept of a fundamental value would apply in an environment where everyone who is in possession of the same *fundamental information* forms identical, *homogeneous expectations* of future share value. This commonly accepted share value would then be the stock's equilibrium price, and the price of shares need not be discovered in the marketplace.

Homogeneous expectations, for good reason, are commonly assumed in academic modeling (to wit, it is a key assumption in the capital asset pricing model). The reason for this assumption is completely understandable—as a simplifying device, assuming homogeneity can make a complex theoretical model tractable. In actual markets, however, expectations are not homogeneous. Rather, participants in possession of identical information concerning a company's fundamentals generally form *divergent expectations* based on that information. A divergence in beliefs is attributable to the sheer magnitude, complexity, incompleteness, and possible unreliability of the information set that pertains to a specific company, an industry, or the broad economy. Simply stated, in a divergent expectations environment, if some participants think, for instance, that a stock should be valued at $25 a share while others assign a value of $30, what is the stock's fundamental value, $25 or $30? The answer is "neither." Stocks cannot have fundamental values when the expectations of market participants are divergent. In a divergent expectations environment, share prices are not discovered in the research offices of the analysts—they can be found only in the marketplace where buy and sell orders meet and are turned into trades.

4.1.1.2 Public Goods

The ability of an exchange to deliver reasonably accurate price discovery is of overriding importance. It is not just the parties to a trade who care about price; a far broader public uses exchange-produced prices for a wide spectrum of purposes that include marking to market, derivative pricing, valuations of mutual fund cash flows, estate valuations, and dark pool pricing. To turn to a nautical analogy, an exchange-produced price shares properties in common with a lighthouse. A lighthouse aluminates the presence of a harbor or the location of a rock; an exchange-produced price sheds light on the value of shares. The beam from the lighthouse benefits any ship that is passing in the night; the light cast by an exchange-produced price benefits the broad investment community.

In economic terms, both a lighthouse and an exchange produce a *public good*. It is well understood in economics that public goods are undersupplied in a private economy and, accordingly, that they must be provided by government. This is indeed the case for a lighthouse, and it is for an exchange as well. Regarding price discovery, an exchange performs a quasi-governmental function of major importance.

Another key consideration is *liquidity* provision. Price discovery and liquidity provision interact in a mutually supporting manner: one would expect price discovery to be sharper in a more liquid market and, reciprocally, that liquidity provision would be more forthcoming in a market that delivers better price discovery. Liquidity, however, is a slippery concept to define and hard to measure; the accuracy of price discovery is even more difficult to quantify (as we have said, *equilibrium values* are not observable).

The quality of price discovery is assessable, however. For one critical reason, this can be done with the use of an intraday volatility metric. The reason? Prices, in searching for equilibrium values, exhibit accentuated volatility. Here is how it works.

4.1.2 Mean Reversion, Returns Autocorrelation, and Accentuated Volatility

The price path from one equilibrium to another rarely follows a straight line. Rather, prices bounce around, describing a jagged path that, with momentum moves (and herding), can cause prices to overshoot new equilibrium values and then reverse course. Prices that systematically fall after having risen (or which rise after having fallen) are said to *mean revert*. A good way to visualize mean reversion is to picture prices first swinging up and then down (or down followed by up) within a trading range. When prices mean revert, a sequence of returns (price changes) is negatively autocorrelated. With negatively autocorrelated returns, prices are not following a random walk. Instead, price increases (or a run of increases) are more apt to be followed by decreases, and price decreases (or a run of decreases) are more apt to be followed by increases.

Mean reversion and its counterpart, negative returns autocorrelation, are present in short-period price movements (e.g., intraday returns), but they decay as one moves to returns measured over longer intervals of time (e.g., a day or more). The price volatility accentuation that is associated with negative returns autocorrelations also decays as one moves to longer measurement intervals. Consequently, the quality of price discovery can be inferred by matching very-short-period price volatility with longer period price volatility.

In a *frictionless* world of perfectly accurate price discovery (that is, in a random walk world), the variance of returns will increase proportionately with the length of the interval used to measure them. For instance, the variance of a distribution of five-day returns will be five times that of a distribution of one-day returns. Thus, a five-day returns variance that is *less* than five times a one-day returns variance indicates that the one-day returns are negatively autocorrelated (i.e., are mean reverting). Equivalently stated, the lower five-day variance is evidence that the one-day return variance is *accentuated* (not that the five-day return variance is depressed). We suggest, first and foremost, that the accentuation is attributable to price discovery being a complex, noisy process which is replete with jagged price moves, overshooting, and mean reversion.

For this reason, the quality of price discovery can be inferred from an intraday volatility analysis. To do so, alternative volatility measures can be employed, with the most popular being variance (or standard deviation) and a high-low range.

4.1.2.1 Volatility Analysis: Evidence

Alan and Schwartz (2013) assessed the level of intraday volatility for a sample of 30 Dow stocks, presenting examples of stock/day-specific opening half-hour volatility for the year 2011. In this subsection, we present a condensed version of the relevant part of that paper.[1]

The purpose of Alan and Schwartz's analysis was not to assess an average level of volatility across a large, all-inclusive set of stocks, but to hone in on the higher levels that volatility can reach in a brief, opening half-hour interval. To achieve this, for all US stocks for each trading day in 2011, they first calculated, for each stock on each day, an opening volatility measure and a spread-adjusted opening volatility measure that are based, not on a variance statistic, but on a stock's high-low price range:

$$\text{Opening volatility} = \frac{P^{\text{max}} - P^{\text{min}}}{P^{\text{mean}}} \tag{4.1}$$

$$\text{Adjusted volatility} = \frac{P^{\text{max}} - P^{\text{min}} - Spread}{P^{\text{mean}}} \tag{4.2}$$

where P^{max}, P^{min}, and P^{mean} for a given stock and day are the highest, lowest, and average trade prices, respectively, during the first half-hour of trading (9:30 AM to 10:00 AM), and *Spread* is the (time-weighted) average bid-ask spread over the same half-hour interval. The opening high-low volatility measure captures the range of price movements over the 30-min period; to get a sharper read on price discovery, this measure is adjusted by subtracting the bid-ask spread from the interval's high-low prices.

---

[1] This material is printed with permission of the *Journal of Portfolio Management*.

Table 4.1 Selected stock/day examples of opening volatility

Company Name (Ticker)	Date	Avg Price	Hi-Lo	Spread	Volatility	Adjusted Volatility	Group*
JOHNSON & JOHNSON (JNJ)	04/06/11	$59.82	$0.20	$0.01	0.33%	0.31%	5
BOEING (BA)	06/29/11	$72.34	$0.38	$0.03	0.53%	0.49%	6
HOME DEPOT (HD)	04/20/11	$38.34	$0.25	$0.01	0.65%	0.62%	7
MERCK (MRK)	09/27/11	$32.06	$0.25	$0.01	0.78%	0.74%	8
TRAVELERS COMPANIES (TRV)	08/04/11	$53.32	$0.48	$0.02	0.90%	0.86%	9
EXXON MOBIL (XOM)	05/26/11	$81.92	$0.82	$0.01	1.00%	0.99%	10
PROCTER & GAMBLE (PG)	04/12/11	$62.52	$0.71	$0.01	1.14%	1.12%	11
MCDONALDS (MCD)	10/05/11	$86.19	$1.12	$0.04	1.30%	1.26%	12
WALMART (WMT)	01/20/11	$55.66	$0.80	$0.01	1.44%	1.41%	13
AMERICAN EXPRESS (AXP)	09/26/11	$46.59	$0.77	$0.03	1.65%	1.60%	14
UNITED TECHNOLOGIES (UTX)	02/24/11	$82.74	$1.53	$0.03	1.85%	1.81%	15
UNITEDHEALTH GROUP (UNH)	12/08/11	$49.25	$1.05	$0.02	2.13%	2.08%	16
VERIZON (VZ)	08/01/11	$35.80	$0.89	$0.01	2.49%	2.46%	17
DU PONT (DD)	08/05/11	$47.88	$1.45	$0.02	3.03%	2.99%	18
JPMORGAN CHASE (JPM)	08/25/11	$37.69	$1.57	$0.01	4.17%	4.14%	19
DISNEY (DIS)	08/10/11	$30.34	$2.31	$0.02	7.61%	7.55%	20

(*There are no Dow stock observations in the first four groups, therefore our table starts from Group 5)

The stocks were next sorted by their adjusted volatility and divided into 20 groups of equal numbers. Group 1 comprised the stock/day observations with the lowest adjusted volatility, and Group 20 comprised the stock/day observations with the highest adjusted volatility.[2] From this all-inclusive set of stocks, the Dow stocks only were selected for the analysis. For each of the 20 groups, Alan and Schwartz selected the single Dow stock that had the highest single-day volatility in the group. The process resulted in 16 observations, which are shown in Table 4.1. Note that no Dow stock/day observation was located in any of the four lowest volatility groups.

Table 4.1 gives the company name and ticker, date of the observation, average price during the opening half-hour, dollar difference between the highest and the lowest price, average spread, opening volatility, spread-adjusted opening volatility, and group to which the observation belongs. On the low end of the spectrum, on April 6, 2011, Johnson & Johnson (at the time, a $60 stock) had a $0.20 price fluctuation in the first half-hour, a spread of $0.01 (2 basis points), and an adjusted volatility of 0.31 %. At the high end of the spectrum, Disney (at the time, a $30 stock), on August 10, experienced a $2.31 price fluctuation in the first half-hour of trading with an average spread of 2 cents (7 basis points). Concurrently, Disney's adjusted high-low was a very substantial 7.55 %. For all 16 observations, the spread-adjusted price volatility displayed in Table 4.1 is indicative of a component of volatility that we suggest represents appreciable price discovery noise.

[2] Alan and Schwartz further imposed a price filter that restricts the sample to stocks in the $30–$100 price range.

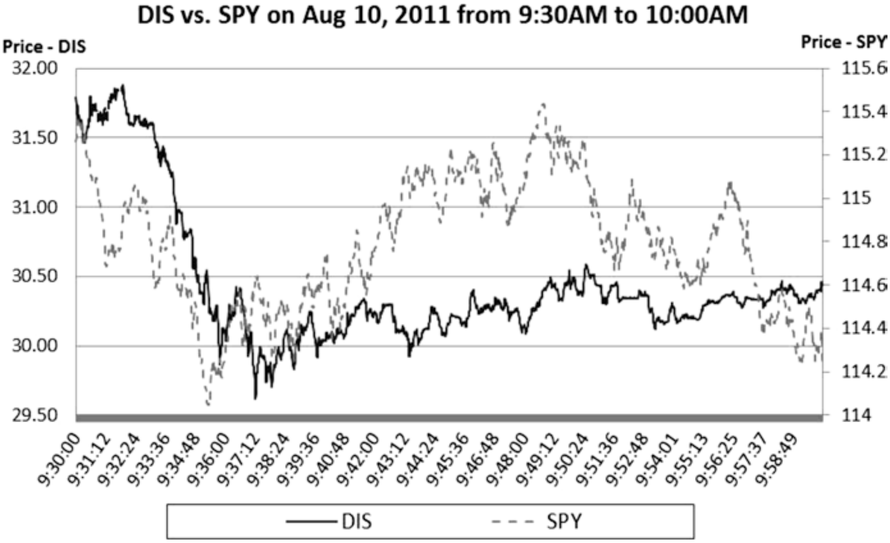


Fig. 4.1 Price path for SPY vs. DIS during the opening half-hour interval on August 10, 2011

Of particular interest is the high end of the spectrum. On this end, Disney, on August 10, 2011, clearly stands out; let us focus on it. On that date, the market for Disney (and the broad market as well) was under stress, as the markets were profoundly rattled by the European debt crisis. Might this explain the high first half-hour volatility? Macro uncertainty is certainly an underlying causal factor, precisely because price discovery is more difficult at times when uncertainty is high and people's expectations about what the future will bring are more divergent. Nevertheless, the question remains this: what could account for one person buying shares at a price that was 7.61 % percent higher than the price at which someone else sold shares within the same half-hour interval when the average spread was only $.02, as occurred on August 10 for Disney stock?

Alan and Schwartz questioned whether or not a fresh news release from Europe (or any other news event during that particular half-hour) could be the cause. A search of LexisNexis revealed no major news announcements at this time for either Disney or the broad market. Neither does Disney's price path suggest the advent of a major news announcement in the opening 30 min of trading on that day. Figure 4.1 shows, second by second, for that first half-hour, how Disney's (DIS) price evolved, side by side with the price of the SPDR S&P 500 ETF (SPY).[3] In Fig. 4.1, DIS's prices are on the left-hand axis and its chart is the solid line; SPY's prices are on the right-hand axis and its chart is the dashed line.

[3] To suppress the effect of price changes attributable to the bid-ask spread and to reduce the effect of out-of-sequence reporting, the prices shown in the exhibit are averages for all trades that occurred in each of the 1800 s that comprise the first half-hour (on that day, DIS averaged 37 trades per second, while SPY averaged 124 trades per second).

For DIS, there is initial volatility and a bump up in the first minute, a predominantly downward trend until 9:37 AM, a predominantly upward trend until 9:50 AM, falling prices for the next couple of minutes, and lastly an uptrend to 10:00. The picture for SPY is simpler: falling prices until 9:35, an upward trend until 9:48, and primarily falling prices to 10:00. Comprehensively viewed, both paths display mixtures of trending and reversals, and the two paths are weakly correlated with each other (the correlation is .19 for 30-s returns and .47 for 1 min returns).

From this evidence, one can infer that intra-half-hour news release is not the cause of the observed price movements for DIS. We suggest that the more plausible cause is the dynamic process of price discovery. Apparently, the August 10 opening price did not adequately reflect the broad market's desire to hold Disney shares. We suggest that the substantial price changes which ensued for at least the next 30 min largely reflected the market searching for a price that better balanced the opposing pressures exerted by a diverse population of buyers and sellers whose expectations, given the greater uncertainties that prevailed at that time, were on that day strikingly divergent.

After having focused on one stock (DIS) in particular, Alan and Schwartz proceeded to consider the full set of 30 Dow stocks over all 252 trading days in 2011. In this assessment, each stock/day observation was assigned to a volatility group.[4] Summary statistics of the adjusted opening volatility for each of these groups are given in Table 4.2. The mean, adjusted volatility ranges from 0.28 % for the lowest volatility group to 5.56 % for the highest volatility group.[5] The faster rise in average volatility among the higher volatility groups is striking: while group 18 has an average volatility of 2.71 %, the average reaches 3.49 % in group 19, and 5.56 % in group 20. Table 4.2 also shows the number (N) and the percent (%N) of Dow observations in each of the 20 groups. Out of the volatility observations for all Dow stocks in 2011, roughly 43 % fall into groups 11–20. In other words, almost half of the Dow stocks experienced an opening volatility that is higher than the median volatility across all stocks. Clearly, it is not just the small cap stocks that experience high volatility—the largest stocks in the economy clearly exhibit accentuated volatility in the first half-hour of trading as well.

4.1.2.2 Monitoring Volatility

Having a volatility auction at times of high volatility insures having (1) a price discovery process in place even when continuous trading has to be interrupted due to larger price movements, (2) and not only allows for the pricing of the underlying stock, but also provides a price point for the respective derivatives instruments related to that stock. Related to indexes, the calculation of their values can (3) continue and is possible at any time during normal trading hours and, like for stocks, (4) any derivative product defined based on such an index can continue to be priced and traded (Fig. 4.2).

---

[4] The same stock was allowed to fall into different volatility groups on different days.

[5] Except for the three highest volatility groups, means and medians are virtually identical.

Table 4.2 Summary statistics of the adjusted opening volatility by group

Group*	Mean	Median	Min	Max	N**	% N
5	0.28%	0.28%	0.16%	0.31%	77	1.02%
6	0.42%	0.42%	0.32%	0.49%	607	8.03%
7	0.56%	0.56%	0.49%	0.62%	939	12.42%
8	0.68%	0.68%	0.62%	0.74%	1032	13.65%
9	0.80%	0.80%	0.75%	0.86%	854	11.30%
10	0.92%	0.92%	0.87%	0.99%	780	10.32%
11	1.05%	1.05%	0.99%	1.12%	754	9.97%
12	1.18%	1.18%	1.12%	1.26%	600	7.94%
13	1.33%	1.33%	1.26%	1.42%	503	6.65%
14	1.50%	1.49%	1.42%	1.60%	412	5.45%
15	1.70%	1.70%	1.60%	1.81%	294	3.89%
16	1.94%	1.93%	1.81%	2.09%	259	3.43%
17	2.25%	2.24%	2.09%	2.46%	197	2.61%
18	2.71%	2.68%	2.46%	3.02%	148	1.96%
19	3.49%	3.40%	3.03%	4.14%	78	1.03%
20	5.56%	4.86%	4.19%	8.89%	26	0.34%

*There are no Dow stock observations in the first four groups, therefore our table starts from Group 5.

**Total number of observations is 7,560 (30 stocks * 252 trading days).

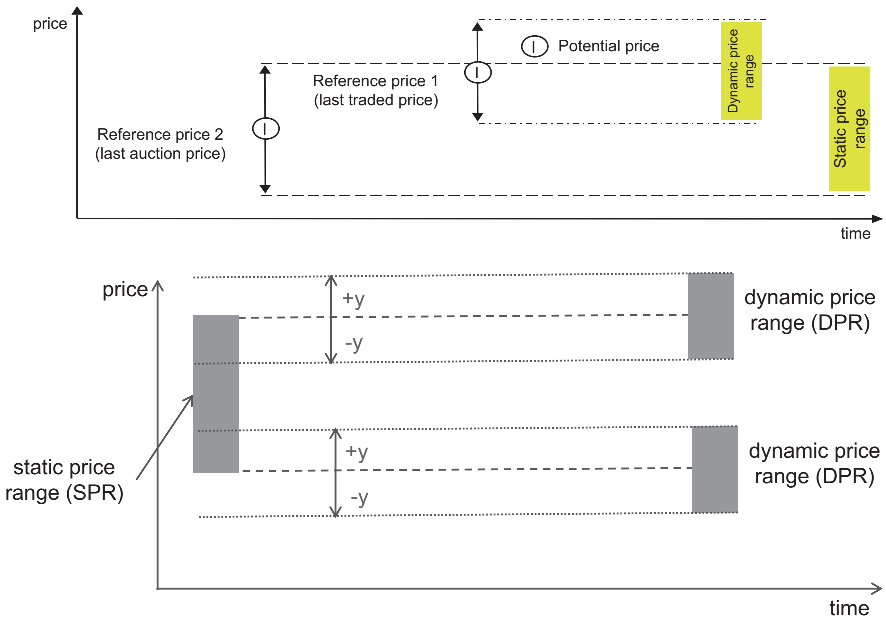


Fig. 4.2 Dynamic vs. static price range

An example illustrating volatility is shown in Fig. 4.3:

1. After rumors emerged in the market that CME Group was planning to make Deutsche Börse Group an acquisition offer, DBG's share price rose to a maximum of € 52.30 (+12 %).
2. After the communication of an ad hoc announcement, the price dropped to € 47.50. The share closed at € 49.30 (+5.6 %) with a turnover surpassing the daily average on 25 February three times.

Trading volume slowly decreased as DBG's communication department denied rumors until the release of the ad hoc (Fig. 4.4).

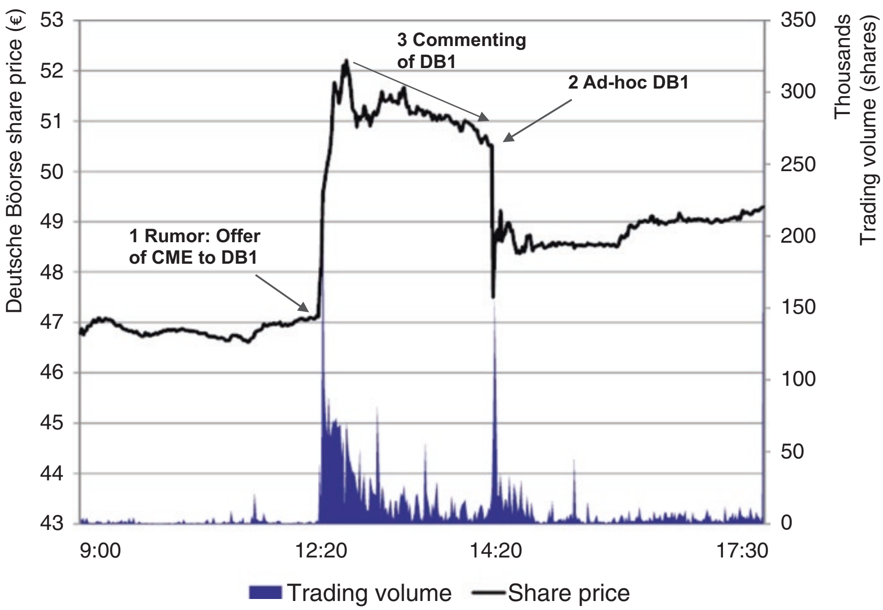


Fig. 4.3 Intraday volatility (Deutsche Börse Group example, 25 February 2013)

Volatility Interruptions in DB1 Shares on 25 February 2013						
No.	Price before VI	Start of VI	End of VI	Price after VI	Volume	Comments
1	48.120 €	12:18:02	12:20:25	49.000 €	29,025	
2	50.440 €	12:25:54	12:28:17	50.600 €	26,469	
3	52.110 €	12:36:33	12:38:56	51.900 €	5,208	
4	50.950 €	14:19:17	14:22:07	48.000 €	46,821	Extended Volatility Interruption
5	47.500 €	14:22:07	14:24:30	47.995 €	78,951	

Fig. 4.4 Volatility interruptions in DB1 Shares on 25 February 2013

4.2 Market Structure

Having recognized (a) the importance of an exchange delivering reasonably accurate price discovery, (b) that the quality of price discovery can be assessed by an intraday volatility metric, and (c) that intraday, first half-hour volatility can be strikingly high, we have one further matter to address at this time: the relationship between the quality of price discovery and a market's architecture.

By market architecture, we are referring to an exchange's rule book, trading systems, and technology. With each of these, clear alternatives exist. Here are some highlights. In a continuous market, trades that are generally bilateral are made whenever a buy and sell order meet or cross in price; in a periodic call market, orders are batched together for simultaneous execution at a single point in time at a single price. On an organized exchange, price is the primary rule of order execution (highest bids are matched with lowest offers), but when the most aggressive orders are tied in price, a secondary rule of order execution is called for; the rule could be time priority (first in, first out), size priority (the largest orders execute first), or pro-rata execution. Designated *market makers* may or may not be included to facilitate liquidity supply. Along with standard market and *limit orders*, an assortment of alternative order types and instructions are generally available (e.g., fill-or-kill orders, all-or-nothing orders, and hidden or iceberg orders). Small retail orders are typically handled in one way, and large block orders in another. A marketplace may be integrated or fragmented, and trading can be transparent or opaque. Systems can be predominantly electronic or driven by human intermediaries. A trading environment may be based on a single modality or it can be a hybrid.

This overview of market structures can be extended in both scope and detail, and we turn to major alternatives later in this chapter. The important point to make at this time is that market quality is not an exogenous variable it very much depends on how order flow is integrated in the process of delivering trades and producing prices. Choice exists, and unanswered questions as to what is best persist. Market architecture remains a work in process.

4.2.1 Continuous Trading in Order-Driven Markets

4.2.1.1 The Link to Continuous Trading: The Spread

No spread exists as a *call auction* book builds with buy orders meeting and crossing sell orders in price. But, after the call has been completed, unexecuted orders remain on the book (unless otherwise instructed) and, because all matching and crossing orders have been executed and are no longer on the book, a spread necessarily exists between the highest posted bid and the lowest posted offer in the continuous market that follows the call. This spread between posted orders continues to exist as the continuous market progresses, widening with the elimination of previously posted orders and shrinking with the arrival of new orders that are not priced aggressively enough to execute upon arrival. The book at the completion of a call, and the spread between unexecuted orders that characterizes the start of the continuous market, is shown diagrammatically in Fig. 4.5.

After applying the matching rules, the market in the above traded stocks looks as follows:

- The execution price (EP) is 100: three shares are matched on each side and both orders execute fully.
- The spread is two, which means for the investor: the best bid to sell to is 99, the best offer to buy at is 101. If one were to inquire what the market is, the answer would be "99–101."
- The remaining orders in the *consolidated limit order book* (CLOB) stay there, ready for matching if incoming orders drive the price in their direction and a match is reached. These booked orders build both the market's breadth (at the price at which they have been placed) and its overall depth, thereby making the market more liquid.

Because there is no surplus of unexecuted orders at the execution price of 100, no order is left on either side of the book at 100. Therefore, the spread is now 99–101.

Referring again to Fig. 4.5, if two orders rather than one were placed at 100 (=EP), the cumulated number of buy orders at 100 would be four. Hence, one order would be left on the book on the buy side after the call has been completed. Accordingly, the quotes and their attending spread after the execution of three orders on each side would be 99–100 and 1, respectively. Reciprocally, if the overhang of one order was on the sell side, the quotes and the spread at the open of the continuous market would be 100–101 and 1, respectively.

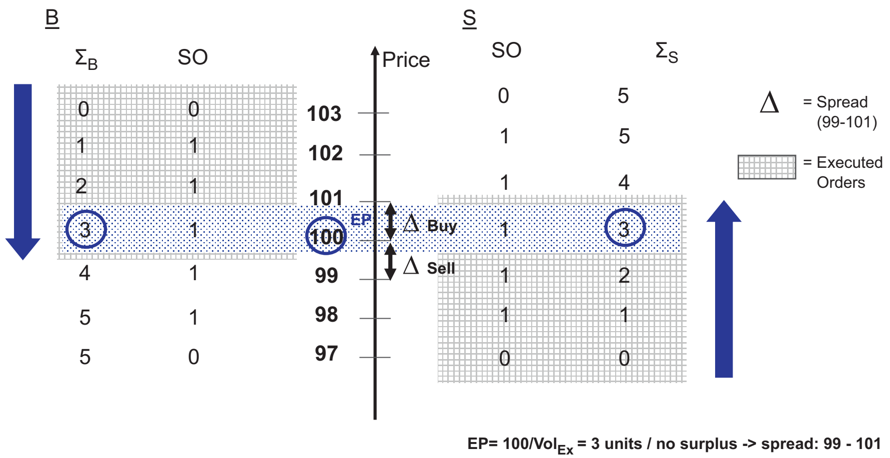


Fig. 4.5 The origin of the spread

Market Orders	Limit Orders
• Execute at the best counterpart • Execute immediately • Immediacy/liquidity demanding	• Execute at price of the order • Delayed or no execution • Immediacy/liquidity supplying

Fig. 4.6 Market orders vs. limit orders

Several broad points can be made about order placement in a continuous market environment[6]:

- Competition among dealers and limit order traders keeps spreads tighter.
- Dealer spreads depend on the costs they incur when providing immediacy.
- The threat of informed trading widens spreads.
- Uninformed traders in particular bear the cost of paying spreads that are wider because of the risk of adverse selection.
- Spreads are tighter for more liquid, better known, large cap stocks.
- Large anonymous traders are widely thought to be better informed.
- Limit-order traders give options to traders who can respond more quickly to changing market conditions.

4.2.1.2 Market Orders and Limit Orders

We alluded to some of the alternatives for market structure in the previous subsection. The two basic structures are continuous order-driven markets and continuous quote-driven markets, where "continuous" means that a trade can be made at any point in time that the market is open and a buy and sell order either meets or crosses in price. In a continuous market, trades are generally bilateral (i.e., one trader's buy order executes against another trader's sell order) as distinct from a call auction where trading is generally a multilateral (batched) matching (Fig. 4.6).

In a pure order-driven market, the orders of some public traders set the prices at which other public traders can buy or sell without the participation of an intermediary. In a pure quote-driven, dealer-intermediated market, a dealer's ask quote establishes the price at which a public trader can buy shares, and a dealer's bid quote establishes the price at which a public trader can sell shares. In this section, we focus on the former, the continuous order-driven market.

The viability of an order-driven market depends on the willingness of some public participants to place limit orders, and on the willingness of other public participants to place market orders. Without limit orders, market orders could not

[6] Schwartz, Robert A./Francioni, Reto (2004): Equity Markets in Action. The Fundamentals of Liquidity, Market Structure & Trading. Hoboken: Wiley.

execute; without market orders, limit orders could not execute. They need each other. Without market orders and limit orders coexisting, an order-driven market would fail. In essence, the limit order placers provide liquidity and immediacy to market order traders who are seeking liquidity and immediacy.

A limit order is so named because the trader who has placed it has stated a price limit at which shares are to be bought or sold. For a buy limit order, the price limit is the maximum share price the trader is willing to pay; for a sell limit order, the price limit is the minimum share price the trader is willing to accept. Alternatively stated, at any price greater than the buyer's limit, the buyer does not wish to acquire the shares; at any price lower than the seller's limit, the seller does not wish to dispose of the shares.

Market orders, on the other hand, are unpriced orders. The trader who has submitted a market buy order is willing to buy at the lowest posted offer, which would be the price established by the most aggressive (lowest priced) limit sell order. The trader who has submitted a market sell order is willing to sell at the highest posted bid, which would be the price established by the most aggressive (highest priced) limit buy order.

Limit orders are generally posted on the limit order book, with buy limits placed below the lowest market ask and sell limits placed above the highest market bid. A limit buy order priced above the best market ask (or a limit sell order priced below the best market bid) is referred to as a *marketable limit order*. Marketable limit orders that are larger than the best market ask (or bid) will *walk the book* (that is, buy orders will execute at successfully higher prices and sell orders will execute at successfully lower prices) until they are fully executed or reach their limit price, at which point any unexecuted portion of the order will be entered in the book as a regular limit order.

4.2.1.3 Costs and Benefits of Market Orders and Limit Orders

Trading by market order conveys one benefit: it enables the participant to trade immediately and, in so doing, to achieve certainty of execution. But a market order strategy entails a cost. Assuming that the order is not large enough to walk the book, the market order trader buys at the ask or sells at the bid and, in so doing, pays the spread. The spread, however, is the cost of a round trip, and thus half of the spread is taken to be the cost of each leg of a round trip.

Trading by limit order saves the spread, but incurs a cost of its own: a limit order on the book may not execute and, if it does execute, it might do so for an undesirable reason. Limit order traders, like market makers, are posting quotes that enable others to trade. A market maker is in business to provide liquidity and immediacy to others. He will sell, not because he wishes to hold fewer shares, but because a public participant wishes to buy; or he will buy, not because he wishes to hold more shares, but because a public participant wants to sell. A limit-order trader, on the other hand, seeks to buy or to sell for the precise purpose of adjusting his or her portfolio holdings. If he or she posts a limit order and it does not execute, he or she has failed to make that portfolio adjustment and that, to him or her, is a cost.

A limit order executes when it (a) gets to the top of the book, (b) gains time priority by being first in the queue at the best bid or offer, and (c) is hit by a market order. The trade-initiating market order could have been part of a temporary buy/sell imbalance, or it could have been motivated by news that, from the perspective of the limit-order placer, was adverse information. When the limit order executes because of adverse information, the limit-order trader bears the cost of "adverse selection" and suffers what is known as "ex post regret."

But trading by limit order is beneficial when the order is executed because of a temporary buy/sell imbalance. In the microstructure literature, the imbalance is attributed to participants buying and selling shares for their own "liquidity" purposes when a fresh receipt of cash is realized or a new need for cash is incurred. A buy/sell imbalance that is informationless can push price to (and also past) the price of a posted limit order, trigger an execution, and then revert back to its former level. After his or her order has executed, the limit-order trader benefits from this reversion. The price reversion is referred to as *mean reversion*, the tendency of price to move back to its "mean" (i.e., average) value after it has been pushed away. Mean reversion in prices translates into accentuated short-period price volatility, as we have discussed previously.

Recognizing that the compensation for limit-order trading is realized through mean reversion and accentuated price volatility, we note that a certain amount of mean reversion (and the associated volatility accentuation) is a natural property of a continuous, order-driven trading environment. If the limit order book is very thick and the bid-ask spread is tight, price dislocation and mean reversion will be minimal, the compensation for placing a limit order will be low, and fewer of these orders will be placed. At the other end of the spectrum, if the book is thin and spreads wide, mean reversion will be strong and a greater number of limit orders will be placed. When the book is in balance, the spread is just wide enough and mean reversion is just strong enough to appropriately compensate the limit-order traders for accepting the risk of not executing, along with the risk of executing because of adverse information change.

This balance between limit orders and market orders also underlies the natural (an economist would say "equilibrium") size of a stock's bid-ask spread. A spread that is "too tight" reduces the benefit of trading by limit order but does little to reduce the risks of non-execution and adverse selection. Thus, a spread that is "too tight" leads to more market orders being placed relative to limit orders, and hence the spread is widened. At the other end of the spectrum, a spread that is "too wide" leads to more limit orders being placed relative to market orders, and hence the spread is tightened. When the spread is of appropriate magnitude, the likelihood of it widening when it next changes equals the probability of it tightening.

We conclude this discussion with the thought that, because the risks of trading by limit order cannot be eliminated by placing a limit order sufficiently close to a counterpart quote that has already been placed on the book, a bid-ask spread is a natural property of a continuous order-driven market, just as is mean reversion and accentuated short-period price volatility.[7]

4.2.1.4 Transparency and a Consolidated Limit Order Book

It follows from our prior discussion that a participant in a continuous order-driven market is led to make strategic decisions: most importantly, whether to submit a market order or a limit order and, if a limit order, the price at which that order is placed. Knowledge of the configuration of the limit order book is critical to making these strategic decisions, as is information concerning recent prices, quotes, and trades. Simply put, reasonable pre-trade transparency and post-trade transparency are both essential for a continuous order-driven market to operate efficiently.

Order flow consolidation is also important. Consolidating the order flow facilitates enforcing price priority across all orders that have been sent to the market. It also enables a secondary priority rule to be imposed across all orders (e.g., time priority). Consolidated order flow, with price and time priority enforced, bolsters competition between all orders that have been sent to the market. Additionally, the consolidation of market information facilitates the formulation of order placement strategies.

4.2.1.5 Limitations of the Continuous, Order-Driven Market

An order-driven market is an ecology that comprises a variety of participants who interact in a variety of ways: some are buyers, and others are sellers. Some are seeking to trade because of new information, others because of their individual reassessments of share value, and others in response to their personal liquidity needs and cash flows. Some seek to trade by limit orders and others by market order. Some are longer term investors and others are shorter term traders. Some are proprietary traders and others are intermediaries. Some are large, institutional players and others are relatively small retail customers. And so forth and so on. The important point is, for the order-driven market to work efficiently, it must be in ecological balance. If it isn't, the order-driven market can collapse.

First and foremost, for a continuous order-driven market to be viable, it must receive sufficient order flow. If the order flow is inadequate, the possible gains from mean reversion will be insufficient to compensate a sufficient number of traders for placing limit orders (and, by so doing, accepting non-execution risk and adverse selection risk). Hence, the limit order book will be unduly thin. A sparse book and

[7] For further discussion, see Kalman Cohen, Steven Maier, Robert Scwartz, and David Whitcomb, "Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread," *The Journal of Political Economy*, April 1981, pp. 287–305.

a correspondingly wide bid-ask spread impose a cost that discourages the placement of market orders and this, in turn, lowers the probability of a limit order executing. A vicious cycle can develop that results in market failure. For this reason, an alternative to the continuous, order-driven market is generally turned to for smaller cap, less frequently traded issues. A more appropriate market structure for the thinner traded securities is the quote-driven, dealer market. Call auctions can also be profitably employed.

While transparency is an important feature of a continuous, limit order book market, large traders, to contain their market impact costs, seek opacity for their orders. It is clearly inappropriate for a large participant to submit a large block as a market order it would walk the book and, if large enough, could clear out the entire contra-side of the book. Neither would a large block be posted as a limit order transparency would be totally lost and, given its size, the probability of the large order executing completely would be relatively low. Consequently, blocks are not submitted as such to the continuous, limit order book market; rather, they are delivered in a succession of small tranches. The "slicing and dicing" takes time, however, and thus immediacy is not supplied in this market environment. Alternatively, the large orders are commonly submitted to an alternative trading system (ATS), many of which are referred to as *dark pools*.

Recently, technology development has brought to light one further complexity for the continuous market: the incredible speed with which orders can be submitted and turned into trades. Accompanying this speed is the ability to measure time with high-frequency precision. In today's markets, time is measured in sub-second intervals, down to nanoseconds and even microseconds.

Fast order submission, trade execution, and information dissemination are clearly desirable; hyper-fast, however, may not be. Because a bilateral trade is made any time that a buy order and a sell order meet or cross in price, speed is not simply desired in and of itself; in *continuous trading* with supersonic speed, getting to the market quickly is not per se important; getting to the market first is what matters. And it is the race to be first that magnifies the importance of speed in continuous market trading. When sub-second readings matter, the continuous market can become hyper-continuous.

No human can follow the quotes, trades, and prices as they evolve with subsection frequencies. Consequently, high-speed trading decisions are made by computers, and computer-to-computer trading can, at times, lead to some undesirable results (e.g., flash crashes that have been experienced in recent years). The cost of acquiring the technology required to achieve such *high-frequency trading* is enormous. In the HFT world, some participants gain advantages through, for instance, co-location and development of sophisticated trading algorithms.

In a horse race, a winner must if at all possible be declared and, with a nanosecond time clock, winning by a nose can do it. But trading is not simply a horse race. In trading, the sequence of order arrival within tiny, sub-second intervals is not attributable to meaningful, underlying information change, and it conveys little information of fundamental economic importance to other participants.

For the most part, the sequence of order arrival in very brief intervals is a matter of chance, of who has the better technology, and the vagaries of the order flow. Recognizing this, a lot can be said for not adhering to a microsecond time stamp. Alternatively, all orders within, for instance, a 1-s interval could be given the same time stamp, a stamp that identifies the second within which each order has arrived. Thus, multiple orders that arrive in the same second should be given the same time stamp, and may be executed in a single multilateral match with the use of a call auction algorithm to determine the trades and prices.[8] Further understanding this possibility requires knowledge of the call auction approach to trading, a market structure that we turn to in the next section of this chapter.

4.2.2 Call Auctions in Order-Driven Markets

An auction is a standardized procedure for handling and matching orders in a consolidated limit order book for the purpose of establishing a clearing price, the number of shares that will trade at that price, and the specific participants who will participate (and to what extent) in the multilateral transaction. The execution price is the value that maximizes the total number of shares that will trade.

The order book for the call can offer different degrees of transparency:

- Regarding quantity and quality, the book can display information ranging from displaying all orders including price, size, and trader name to simply showing indicated clearing prices but not volume or any other information.
- Regarding time, from seamless to minutes or a couple of minutes.
- Regarding addressee, the professional traders get the market information in real time, interested public parties postponed.
- Regarding data dissemination (which is an additional business for stock exchanges), a variety of combinations of what/when/to whom are in place; here segmentation is key, as well as, e.g., data streams for professional traders or *algo traders*.

The order book is built in the following steps:

1. Buy and sell orders are entered with price and time of entry recorded.
2. Limit orders on each side of the book are cumulated, from the highest price to the lowest for buy orders, and from the lowest price to the highest for sell orders.
3. Market orders are cumulated and included in their respective totals.
4. The cumulated buy orders are matched against the cumulated sell orders.

[8] For further discussion, see Robert Schwartz, "Slow Down, Wall Street," Commentary in *Traders Magazine*, July 2014, and Robert Schwartz and Liuren Wu, "Equity Trading in The Fast Lane: The Staccato Alternative," Invited Editorial, *Journal of Portfolio Management,* Volume 39, Issue 3 Spring 2013, pp. 3–6.

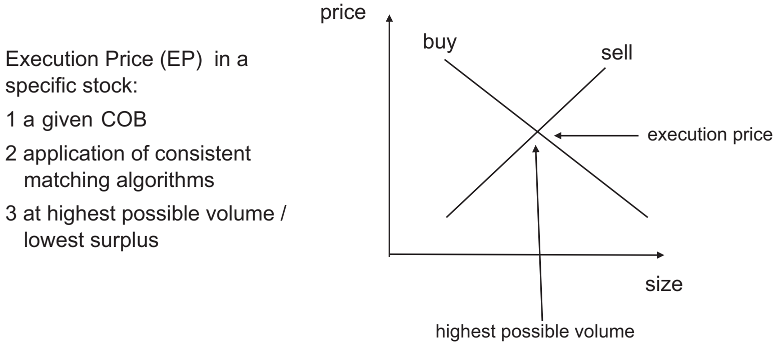


Fig. 4.7 Calculation of the execution price

With this preparation of the central order book, the matching algorithms can be applied to get the execution price (Fig. 4.7). The matching algorithms of an auction are an axiomatic system, whereby the overriding principles are the following:

- For one price without surplus (i.e., a clean cross): most possible executions, maximum turnover.
- For one price with surplus: maximum turnover *and* additional criterion.
- An additional criterion, like market pressure or most recent price, has to be introduced if two prices with the same surplus would be executable.

The algorithmic matching system must fulfil the following criteria:

- Consistency—Equal treatment
- Completeness—No loopholes or gaps in the procedure
- Simplicity—Least possible complexity

The first two bullets are necessary conditions, whereas the third is more of an criterion to improve effectiveness and efficiency.

4.2.2.1 Essentials

There are several necessary preconditions for a functioning call auction.

Fungibility means interchangeability: The shares of a listed company have all the same features and characteristics in content, form, and time. This is why buyers and sellers can trade at exchanges and clear through CCPs, which means netting and offsetting. In order to trade in an auction, the presence of buyers and sellers is mandatory besides a COB. Orders can be stored and deleted in a regulatory, compulsory way and, in applying the matching algorithms, orders get translated into trades. The whole call auction is embedded on the rules and regulations of an exchange (Fig. 4.8).[9]

[9] Cf. paragraph 4.2.2.4.

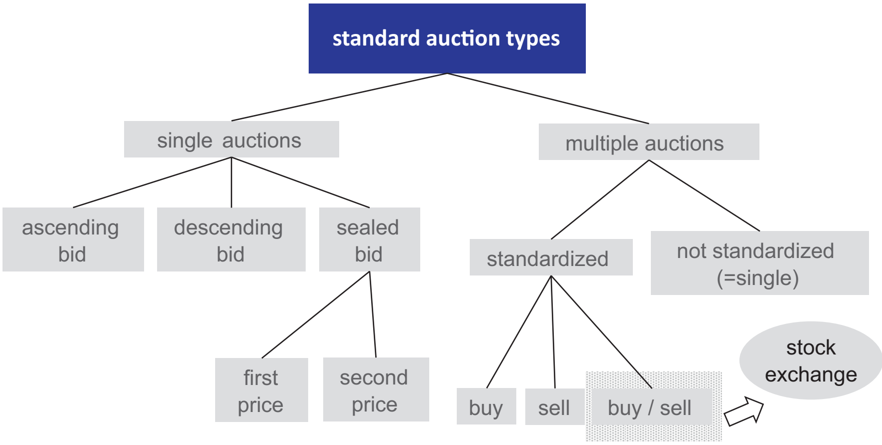


Fig. 4.8 Standard auction types

Auctions can be structured in different ways, as is illustrated in the diagram above. In a single (one-sided) auction, the highest bidder gets the object (e.g., a piece of art). In a descending auction, the first (lowest) bidder gets the object. If sealed bids are allowed, one alternative for the execution price is the first sealed bid; another alternative could be that the second sealed bid is the execution price.[10]

An execution price is determined by every auction, while the sequence of the repeated auctions determines the perfection of continuation (seamless). An auction can also cover the sell *and* the buy side simultaneously, as does an auction at the stock exchange. In contrast to a single auction, this is called a double auction.

4.2.2.2 Double Auction

With a double auction, bids and offers for a specific stock are matched to find the execution price. This price determination is achieved by applying a specific set of rules (matching algorithms).

A double auction can fulfil several functions:

- Open the market (opening auction)
- Reopen the market following a trading halt
- Close the market (closing auction)
- Mimic continuous price discovery by adding several multiple auctions in very short periods of time (even seconds are possible)

Figure 4.9 shows the situations that can occur before an *execution price* (EP) is determined:

[10] This is in accordance with "buy low/sell high," meaning for a single auction: in the first sealed bid it would be the one with the highest price, which is also true for the second sealed bid.

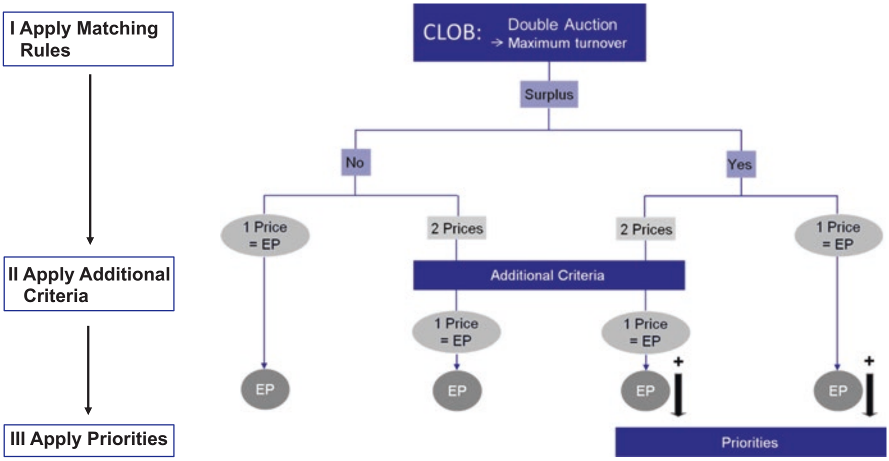


Fig. 4.9 Determination of the execution price in double auctions

If in a double-auction-price-discovery, the right side (*Yes*) applies, two possibilities can occur: either there is one price or there are two prices.

1. In the case of one price, this is the EP. And with an additional priority the handling of the surplus is defined. Some possible priorities are outlined in Sect. 4.5.
2. If two prices are possible, first the criteria to find the EP are applied, and afterwards the priority to handle the surplus.

A graphical outline of a double auction to open the market is shown in Fig. 4.10a, b:

1. Pre-trading: During the pre-trading phase, all incoming orders are collected in the COB. At the same time, orders may also be deleted. The COB is open, and the cumulated breadth and depth of the book can be seen in its entirety.
2. Auction (Deutsche Börse AG example): Once the call auction has started, the book is no longer transparent; there will be only a display of the indicative EP.[11] At a random end, the COB is frozen, so that no order may be entered or taken out of the book. Then the price determination is activated by applying the matching *algorithms*. At the end of the call the EP is defined and the market is cleared. The best (most aggressive) unexecuted orders that remain on the book set the *spread* which applies as continuous trading starts.

Through placing calls within the trading hours, the trading day gets a clear structure (Fig. 4.11).

[11] This is an optional feature.

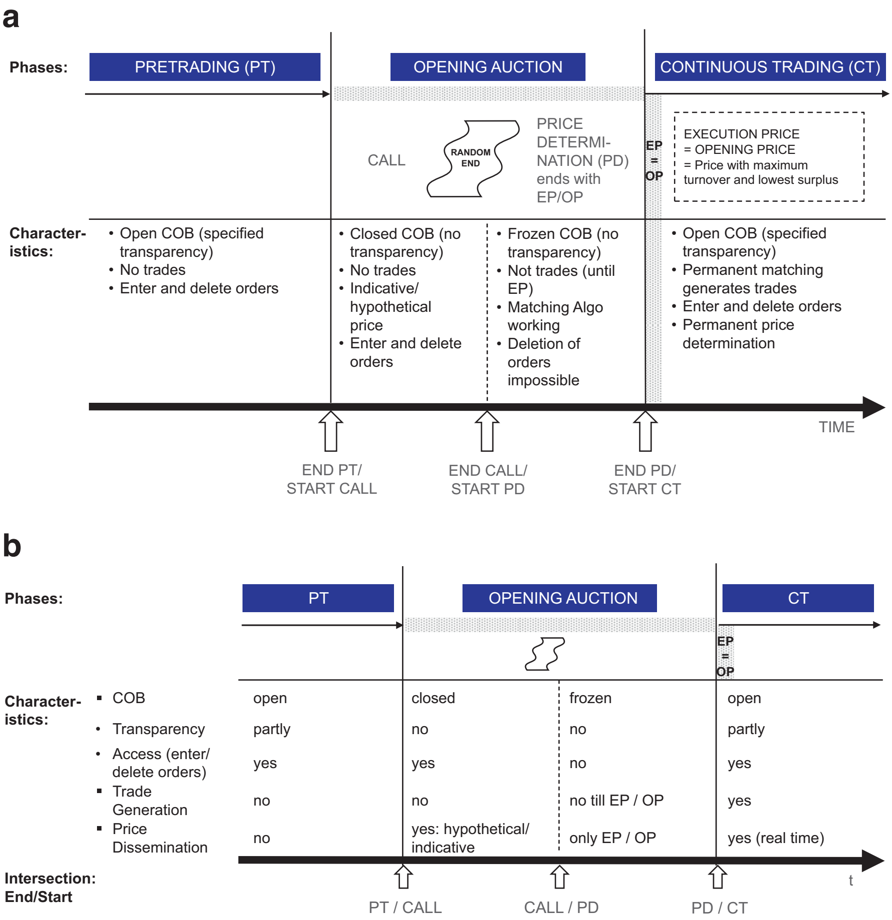


Fig. 4.10 (**a**) Opening the market through double auction. (**b**) Double auction characteristics

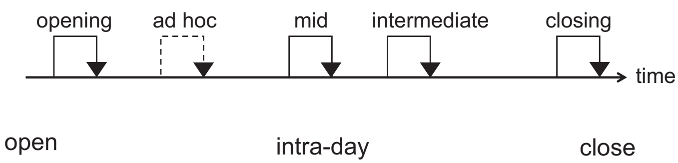


Fig. 4.11 Example structure of a trading day

4.2.2.3 Call Auction

A call auction is an order-driven market. Unlike the continuous order-driven market, in call auction trading orders that could otherwise be matched and executed (in bilateral trading) are held and cleared (in multilateral trading) at a single point in time, and at a single price. The matched and crossing orders which are executed at that single price include buy orders at the call price and higher, and sell orders at the call price and lower.

As noted, the clearing price at a call is determined by selecting the value which maximizes the number of shares that trade. This value is found by matching the cumulated buy orders (cumulating from the highest priced buys to the lowest) with the cumulated sell orders (cumulating from the lowest priced sells to the highest). Because order size and price are not continuous variables, a buy-sell imbalance (surplus) commonly exists at the market clearing price. Surpluses are typically handled by executing orders on the deeper side of the book according to the sequence in which they were transmitted to the market (i.e., by applying a first-in, first-out time priority rule).

During the first price determination in the process of getting an EP during a call auction, three phases have to be distinguished:

In the first phase, the COB has to be built and prepared to start the price determination. Then, in phase two, the actual determination of the EP takes place by applying matching algorithms. And eventually, in phase three, a possible surplus has to be handled.

As noted, if based on the application of the matching rules, two prices are possible, and an additional criterion is necessary to determine the execution price (Fig. 4.12). One of the following three criteria can be applied:

1. *Criterion of smallest surplus*: The objective of the smallest surplus criterion is to minimize the number of unexecuted orders. Therefore, in Fig. 4.13 the execution price (EP) is 99 because, at this price, the surplus is 500 shares while at 98.75 the surplus is 1000.
2. *Criterion of market pressure*: If, after the application of the two above mentioned criteria, two prices are still possible, the third criterion has to be resorted to. For instance, for the two possible prices 99 and 98.75:

   - The trading volume is 3000 shares *and at the same time.*
   - The minimum surplus is 1000 shares.

Because the surplus of the two possible prices is on the buy side, prices are driven up (if both surpluses were on the sell side, prices would be driven down): Therefore, 99 is the execution price (EP).

If the two equal surpluses at 500 shares were on the buy side (at 98.75) and the other one on the sell side (at 99), as shown in Fig. 4.14, the execution price (EP) is 99, because the reference price, usually the last paid price for the preceding day, is 99.50.

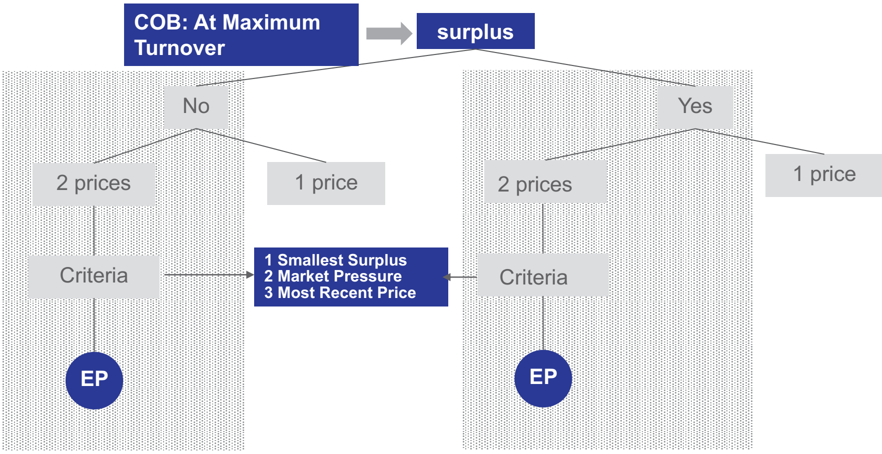


Fig. 4.12 Determination of the execution price in case of two prices

EXHIBIT 4.A2 Setting the Price: The Smallest Surplus Criterion

	Buy Orders, Number of Shares				Sell Orders, Number of Shares			
		Accumulated			Accumulated			
	Separate	Per Price	From Highest Price	Price (Limit)	From Lowest Price	Per Price	Separate	
	400 + 300	700	700	Market				
	....	....	....	> 100				
	200	200	900	100				
	300	300	1,200	99.75				Surplus
Surplus	400	400	1,600	99.50	4,700	500	200 + 300	
	200 + 300	500	2,100	99.25	4,200	700	700	
	800 + 100	900	3,000	99.00	3,500	500	500	500
1000	1,000	1,000	4,000	98.75	3,000	800	100 + 700	
	700 + 200	900	4,900	98.50	2,200	300	300	
				98.25	1,900	300	100 + 200	
				98.00	1,600	100	100	
				< 98	....	....	....	
				Market	1,500	1,500	700 + 800	

Fig. 4.13 The smallest surplus criterion

3. *Most recent price criterion*: In the unlikely case that two EPs with the same surplus but on different side of the market occur, the price that is closest to the last paid price—the reference price[12]—is selected. In the following example of Fig. 4.15, the reference price of 99.60 is closer to 99.0 than to 98.75.

[12] In the rules and regulations of exchange organizations, the reference price is usually the closing price of the previous trading day. The closing price is also used as a reference price for the derivatives market.

EXHIBIT 4.A3 Setting the Price: The Market Surplus Criterion

	Buy Orders, Number of Shares				Sell Orders, Number of Shares		
		Accumulated			Accumulated		
	Separate	Per Price	From Highest Price	Price (Limit)	From Lowest Price	Per Price	Separate
	400 + 300	700	700	Market			
	....	....	....	> 100			
	200	200	900	100			
Surplus	300	300	1,200	99.75			
	400	400	1,600	99.50	4,200	500	200 + 300
	200 + 300	500	2,100	99.25	3,700	700	700
1000	800 + 1,100	1,900	4,000	99.00	3,000	0	0
1000	0	0	4,000	98.75	3,000	300	100 + 200
	700 + 200	900	4,900	98.50	2,700	800	800
				98.25	1,900	300	100 + 200
				98.00	1,600	100	100
				< 98	....	....	....
				Market	1,500	1,500	700 + 800

Fig. 4.14 The smallest surplus criterion

EXHIBIT 4.A4 Setting the Price: The Most Recent Price Criterion

	Buy Orders, Number of Shares				Sell Orders, Number of Shares			
		Accumulated			Accumulated			
	Separate	Per Price	From Highest Price	Price (limit)	From Lowest Price	Per Price	Separate	
	400 + 300	700	700	Market				
	....	....	....	> 100				
	200	200	900	100				
	300	300	1,200	99.75				Surplus
Surplus	400	400	1,600	99.50*	4,700	500	200 + 300	
	200 + 300	500	2,100	99.25	4,200	700	700	
	800 + 100	900	3,000	99.00	3,500	500	500	500
500	500	500	3,500	98.75	3,000	800	100 + 700	
	700 + 200	900	4,400	98.50	2,200	300	300	
				98.25	1,900	300	100 + 200	
				98.00	1,600	100	100	
				< 98	....	....	....	
				Market	1,500	1,500	700 + 800	

*preveious price

Fig. 4.15 The most recent price criterion

If through application of the outlined criteria an execution price is determined and there is no surplus, all orders on both sides of the market are executed at the EP. If there is a surplus, it is calculated as the number of shares at the larger side minus the number of shares at the smaller side. The smaller side, which always executes completely, establishes the number of shares that trade, while shares on the larger size have to be rationed (Fig. 4.16).

Several rationing criteria are equally possible:

Time priority: Time is the most common secondary priority rule (with price being the first priority rule) (Fig. 4.17). The application is as follows:

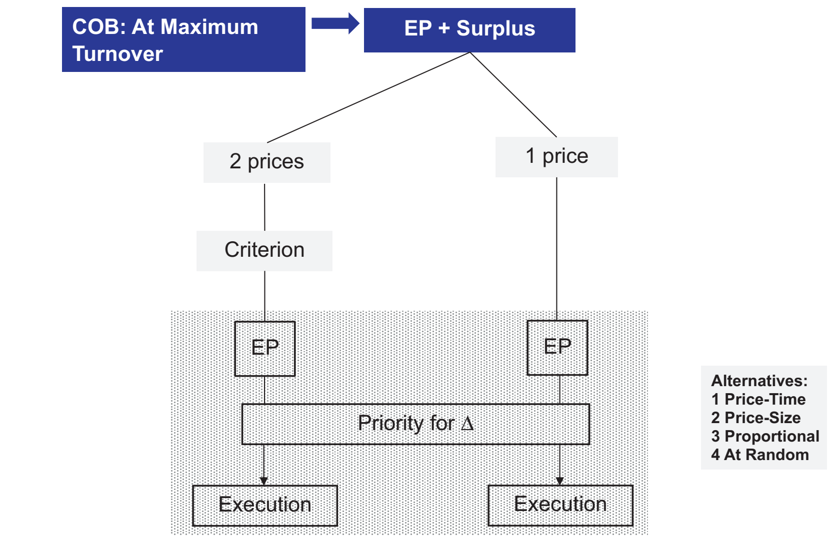


Fig. 4.16 Priorities on how to handle the surplus

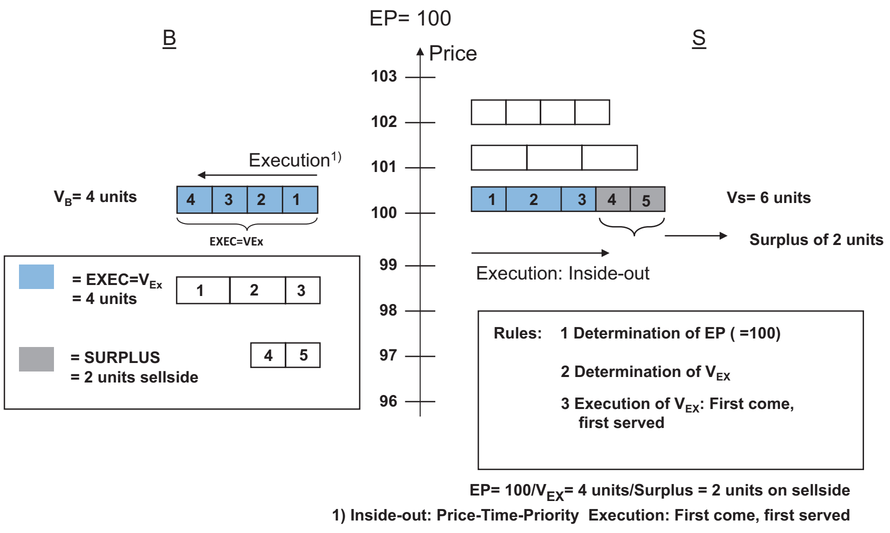


Fig. 4.17 Time priority

1. At the EP (= 100) arrange the orders based on their time stamp.
2. Then apply priority "first come first served/executed":

- On the smaller buy side, all orders are executed: The executable volume at EP is 4 units/shares.

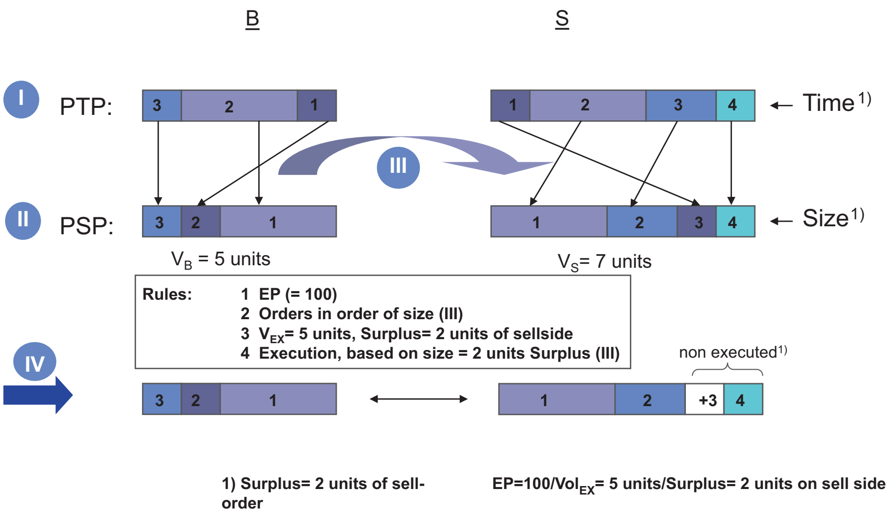


Fig. 4.18 Size priority

- On the bigger sell side, orders 1–3 are executed, and orders 4 and 5 (=2 units/shares) remain in the order book.

Size priority: If at the EP bigger orders receive priority, size priority is applied.

1. The COB is built as follows: two sided, applying price priority.
2. Within one price: size priority (bigger order size before smaller order size).
3. At the EP (=100) the executable volume is the smaller (buy) side with three orders adding up to five shares.
4. Execution:

- The whole buy side is executable: three orders adding up to five shares.
- On the sell side, 5 units—orders 1 and 2—are executable. Orders 3 and 4 are not executable and therefore build the surplus of 2 units or shares (Fig. 4.18).

Proportional-execution-priority: Proportional-execution-priority means that each order of the larger (surplus) side at the EP is executed proportionally to the smaller side (Fig. 4.19). Regarding the example illustrated above this means:

1. The proportion between the smaller (five units) and the bigger side (ten units) is 5:10 which equals 0.5 or 50 %. Therefore,
2. Every order on the bigger side is executed up to 50 %, which means half. The not executed part (five units) remains in the COB.

At random priority: Execution at random means that orders on the larger side are executed randomly until the sum of the executed order reaches the sum of the smaller side (=5 units) (Fig. 4.20). The unexecuted part (=3 units) remains in the COB.

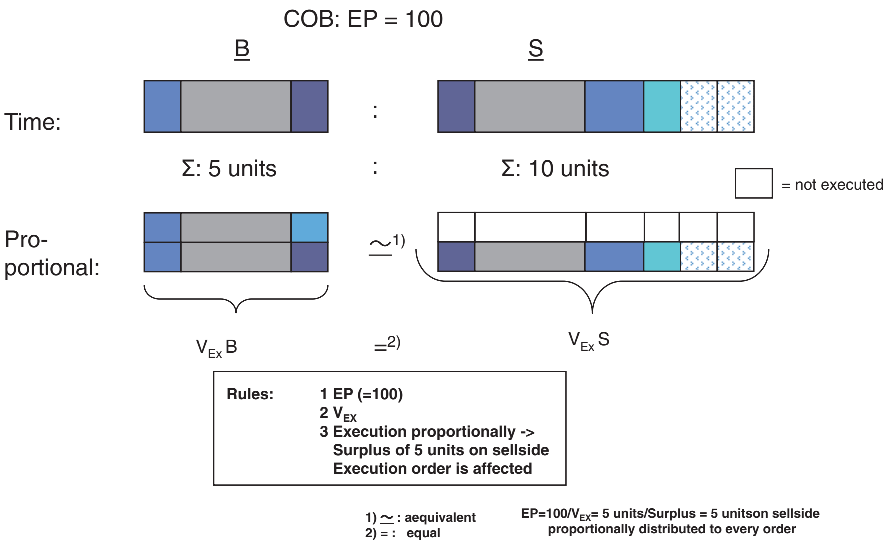


Fig. 4.19 Proportional execution priority

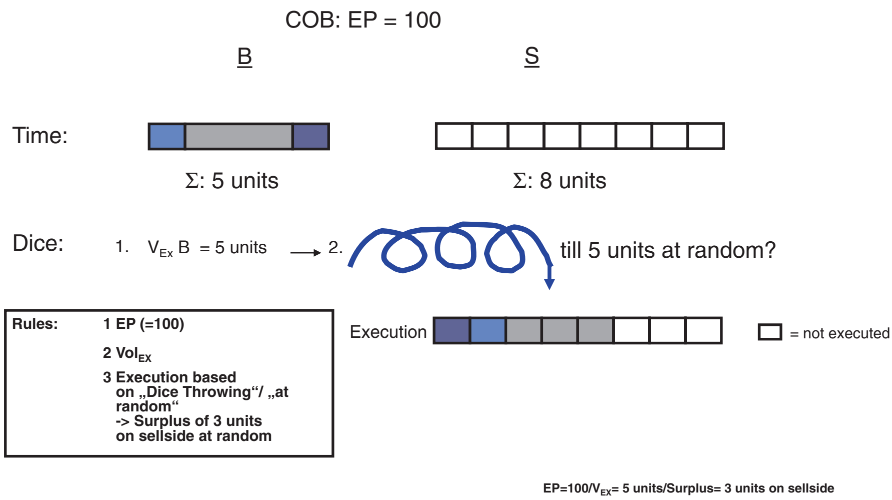
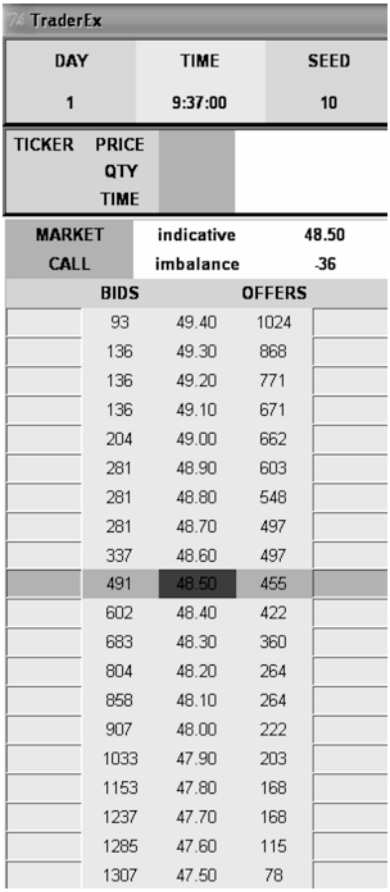


Fig. 4.20 At random priority

Figure 4.21 shows an illustrative limit order book for call auction trading.[13] The middle column displays the prices at which orders have been placed. The column on the left shows the cumulative number of round lot (100 shares) buy orders, cumulating down from the highest price at which each limit buy order has been

[13] Figure 4.21 is a screenshot from TraderEx, a computerized trading simulation software program. For further information, see www.etraderex.com.



BIDS		OFFERS
93	49.40	1024
136	49.30	868
136	49.20	771
136	49.10	671
204	49.00	662
281	48.90	603
281	48.80	548
281	48.70	497
337	48.60	497
491	48.50	455
602	48.40	422
683	48.30	360
804	48.20	264
858	48.10	264
907	48.00	222
1033	47.90	203
1153	47.80	168
1237	47.70	168
1285	47.60	115
1307	47.50	78

Fig. 4.21 Illustrative limit order book for call auction trading

placed, and the column on the right shows the cumulative number of round lot sell orders, cumulating up from the lowest price at which each limit sell order has been placed. Given the displayed array of cumulated buy and sell quantities, 48.50 is the price that maximizes the number of round lots that would trade. At this price, a buy-sell imbalance (surplus) exists: cumulative bids (491) are greater than cumulative offers (455), and the number of round lots that can execute (being the lesser of these two values) is 455. This buy-side imbalance is handled by rationing the buy orders as we have just discussed (the criteria include time priority, size priority, proportional allocation, and random selection).

Note that no other price results in a number of executable round lots greater than 455. One price tick higher, at 48.60, the minimum of the cumulated bids and offers is 337 (resulting in a sell imbalance) and, one price tick lower, at 48.40, the minimum of the cumulated bids and offers is 422 (resulting in a buy imbalance). Thus a price of 48.50 maximizes the number of shares that trade and, accordingly, 48.50 is the clearing price.

Because limit orders submitted to a call execute at the clearing price established for the call, limit orders are price improved (with the exception of those placed at the clearing price exactly). This contrasts with continuous market trading where limit orders on the book execute at the price at which they have been entered. Because limit orders submitted to a call are commonly price improved, they should be priced more aggressively (i.e., higher priced buy limits and lower priced sell limits) than limit orders submitted to a continuous market. Another difference is the muted distinction between market orders and limit orders: market orders submitted to a call auction are nothing other than infinitely aggressively priced limit orders (infinitely high prices for buy orders, and zero prices for sell orders). Moreover, in contrast to continuous market trading, in call auction trading market orders do not execute with immediacy but only when the market is called.

Several advantages attend call auction trading. Batching orders together for point-in-time trading consolidates liquidity temporally. Systematically finding the clearing price with reference to the full set of cumulated buy and sell orders sharpens price discovery. Vis-à-vis continuous market trading, the order batching, single price auction procedure is fairer, and more difficult to manipulate. Recognizing these advantages, one might anticipate that call auctions would be widely used as a trading modality.

Call auctions were prevalent in the early days of trading, but nonelectronic calls had severe shortcomings and, as volumes increased in the precomputer age, the auctions were replaced by continuous trading. However, around the turn of the twenty-first century, calls started to reemerge in markets around the world. They have done so as modern, electronic facilities that are typically being used to open and to close trading in a hybrid combination with continuous trading. As we have noted, call auctions are also used to reopen markets after trading halts.

Uniting call and continuous trading eliminates one disadvantage of a call auction-only model: a participant need not wait for a market to be called in order to trade. The application of computer technology eliminates a second disadvantage of a nonelectronic call: investors can participate in an electronic auction in real time without being physically located on an exchange's trading floor.

When it comes to designing a call auction, a considerable number of alternatives exist. An auction can be totally opaque (closed book) or completely transparent (open book), or it can reveal only partial information about booked orders and an indicated opening price. A secondary trading priority rule (most prevalently time priority) can be applied to the order imbalance at a clearing price only, or to all executable orders on the deeper side of the market. The precise time when a market is called is generally determined by random draw within a prespecified, brief trading interval preceding a preannounced time (e.g., at the open or the close of a trading day). Calls can also be initiated at the request of a participant. A call can accept unpriced market orders, or it can be required that all orders be priced. Call auctions are generally price discovery facilities, but a variant exists: a *crossing network* matches customer buy and sell orders at an exogenously determined price (the midpoint of the bid-ask spread in a concurrently running continuous market, or at the closing price in the continuous market for after-hours trading.

This list of design alternatives can be extended. The important point is that not all calls are alike. As one might expect, some call designs will operate more efficiently than others. Much care must be taken to structure a call properly and, as is always the case with system design, one should recognize that the devil is in the details.

4.2.3 Market Making

4.2.3.1 Quote-Driven Market vs. Order-Driven Market

Order-driven markets consolidate liquidity in a single space—the order book. In the order book, limit orders and market orders representing bids and asks are placed and rules determine how trades occur. Basically, all orders are treated equally. Typically, only the type of order (limit or market), the limit order price, the time of order placement, and the order size (number of shares) matter. All traders can trade with each other in the same way, and there are no specific roles defined or incentives given to perform certain actions. As we have described, liquidity is gathered by limit orders submitted to the order book, before these orders grant other orders the option for an immediate execution.

However, liquidity provision and options to trade might be low in certain market conditions, especially for less frequently traded stocks. In these cases, it can be difficult to sustain continuous trading, and additional sources of liquidity will be necessary.

Market makers as a specific type of intermediary fill this role in many of today's markets. Their role is to provide two-sided markets, which means that they are mandated to continuously post bid and ask quotes to the market, and thus give other market participants the possibility to trade. Those quotes must be good for a minimum size and a maximum spread (the difference between the price of the ask and the price of the bid).

In quote-driven markets, the market is split into liquidity providers and liquidity takers. That split is a main difference between quote-driven markets and order-driven markets. Typically, multiple market makers operate simultaneously as competitors in providing their services to liquidity takers in a marketplace.

All bids (and offers) provided by market makers give other market participants the possibility to sell (and to buy). A market maker's quotes are options to buy or to sell. Liquidity takers cannot trade with each other; they are pure liquidity takers. They have to trade with a market maker.

A trade occurs if and when a market participant chooses one bid or ask of a market maker to trade with. By hitting the bid, or taking the offer, the constituent parts of the trade are determined (price, volume, and the two market participants).

In most dealer-driven markets, there are no secondary priority rules of order execution. Traders can choose the market maker they want to trade with. They can direct their orders to specific dealers, a practice known as "preferencing." Here we see a further difference with order-driven markets which normally do not allow that type of practice.

Multiple market makers can be present in a marketplace. They compete with each other in the provision of liquidity.

4.2.3.2 A Market Maker's Role in an Order-Driven Market

The role of liquidity provision through a market maker can also be attached to an order book in a hybrid trading system. The order book works as described in the section on order-driven trading. In that case, in addition, market makers have the obligation to provide liquidity as in a quote-driven setup. Different from a pure quote-driven setup, they send their quotes into the order book and compete with limit orders in the order book. In this setup, all market participants interact via the limit order book in an equal way. The exclusivity of liquidity provision of market makers is broken up, and market makers' quotes become subject to the matching rules of the order book.

To compensate market makers for conducting their role, they in return receive certain benefits as an incentive. These can be discounts on trading fees or even a suspension from all charges of trading and post-trade clearing. Also, anonymity which is common in today's markets can be abandoned for market makers. Consequently they, and only they, can see with whom they trade. That privilege is supposed to help market makers identify so-called informed traders and thus reduce the market maker's risk of trading with these counterparts.

A further release from a strict quotation requirement is sometimes granted to market makers if they are obliged to provide a quote on request only, and not place it more permanently on the order book. Market participants can request a quote, and market makers must respond by sending the quote into the order book within a defined span of time.

An example of a market maker linked to an order book is the *designated sponsor* on Deutsche Börse's Xetra. The designated sponsor has the obligation to provide a quote, on a constant basis, into the order book of some stocks with low or medium liquidity. In addition, a quote request can be sent and the designated sponsor sees the name of the requestor. Furthermore, discounts on fees are granted to designated sponsors by the market operator.

4.2.3.3 A Market Maker's Role in Low- and Mid-Cap Stocks

Market makers like Deutsche Börse's designated sponsors contract with an issuer to provide their services to the market. The market maker is compensated by the issuer for providing liquidity in his or her stock. The market maker conducts research on this stock and provides analysis to the market. Deutsche Börse measures the performance of the designated sponsors in a stock and publishes performance figures on a regular basis. This information gathering provides important guidance in the process of liquidity provision. Market maker revenues (spread and short-term trading in a mean reverting environment).

Market makers are compensated from two sources for providing their services (cf. "The Equity Trader Course," pp. 243–251): the bid-ask spread and trading the order flow. Market maker trades in a quote-driven market are typically "net trades"; namely a commission is not paid. A market maker realizes the bid-ask spread by buying low and selling high. Competition among market makers leads to a tightening of the spread. Wider spreads increase the market maker's profits, while competition resulting in tighter spreads reduces profits. Market makers with a "*long position*" profit when prices rise, and market makers with "*short positions*" profit when prices fall because they can cover their positions at lower prices.

Market makers need to manage their inventory. By adjusting a quote downward, a market maker attracts buyers who react to his or her aggressive ask. Consequently, inventory goes down. Vice versa, if the quote is raised, the market maker attracts sellers who react to the more attractive posted bid, and the market maker's inventory of shares goes up. To manage inventory, market makers can also trade with each other; this is called "*interdealer trading*."

Revenues may also arise for a market maker when successfully "trading the order flow" (cf. Equity Trader Course, pp. 243–251). If a dealer has a good sense of where the market is going short-term, he or she can profit from this insight. To do so requires the ability to detect trends and mean reverting behavior in the market. Timing is of the essence. A market maker profits when knowing when, on net, to buy or, on net, to sell.

4.2.3.4 Market Maker Costs (Costs of an Unbalanced Inventory and Asymmetric Information)[14]

In both types of markets, the order-driven as well as the quote-driven market, the natural buyers and sellers remain the ultimate source of liquidity. The "naturals" generally seek to hold positions in a portfolio for a longer time. Market makers seek to hold inventories (long or short) on a short time base only. They buy not for their own investment purposes, but to grant others the option to buy or to sell immediately. In so doing, they accept the risk of carrying an undiversified portfolio. Market activity (be it preferencing, volatility of prices, infrequent order flow, and stochastic nature of order flow) makes running an inventory more difficult and costly, and thus increases the spreads a market maker is willing to post.

Market makers, like any other traders, expect to incur losses from trading with better informed market participants. For example: a market maker buying stocks from an informed trader coming in before the stock's price is about to fall will lose from that trade. Market makers are compensated for that loss when

[14] The Equity Trader Course, pp. 248–251.

trading with "liquidity" traders (sometimes also referred to as "uninformed" traders). The volume of dealing with them must be large enough to compensate for trades that a dealer is making with better informed traders. That "ecology" of a quote-driven market is necessary for dealers to stay in business, i.e., for the market to exist.

4.2.3.5 Market Makers as Liquidity Providers[15]

Liquidity provision is the main role of market makers. As noted above, the role involves offsetting temporary imbalances between buyers and sellers (demand and supply) in the market.

In that sense, market making is the immediate provision of liquidity. The market maker is permanently present in the market, supporting liquidity provision on a continuous basis. That concept may be extended to a periodic service when it is linked to a periodic type of trading like the call auction. Market makers attached to a continuous order-driven trading market can also be required to provide liquidity (a quote) to call auction trading. The market maker's presence during the entire call phase in the auction may be required in such a setup.

Whether acting as the single source of liquidity in a "pure" quote-driven environment, or acting in combination with an order-driven format, "hybrid" market makers represent a flexible solution to providing liquidity.

4.2.3.6 Market Makers as Facilitators[16]

Trading only occurs when buy and sell orders meet in both space and time. We have described how market makers temporarily step in when there is an order flow imbalance. They do so by providing two-sided liquidity. Their role can go even further. Their activity may trigger orders which "are not yet" displayed to the market. An active trader may attract more liquidity to come to the market, even in a way that triggers a "burst of trading."

4.3 Functions of Market Models (a Designer's Perspective)

Viewed abstractly, trading is a process of information transformation that produces transactions. The carriers of information are the orders that meet in a market, along with the dialogue conducted by traders. The place for this information exchange is the trading system which comprises either a trading floor where participants meet

[15] The Equity Trader Course, p. 240.

[16] The Equity Trader Course, pp. 240–241, Animation.

face to face, or an electronic system where they meet virtually. The previous section has shown the importance and complexities of price determination, but the challenges of finding a good market structure extend beyond the construction of a robust price discovery mechanism.

In addition to price and quantity, trades comprise information about who trades, the type of asset traded, when and where the trade has taken place, and information about how the trade is settled (i.e., the modalities of the post-trading phase). A chosen market model prescribes how this information is generated from information that has been received. In this sense, a market model is the definition of a function. If we take today's electronic trading systems that have been applied in many market structures, a defined function is implemented through an algorithm that makes an outcome deterministic (i.e., the results are always the same when all of the inputs into the algorithm are identical).

In the text that follows, we briefly discuss the diverse functions that various market models define.

4.3.1 Determination of When a Trade Occurs

For a trade to be triggered, certain conditions must be fulfilled. The market model defines these conditions. Buyers and sellers must be in agreement on all conditions of a trade. Achieving this can be the result of a negotiation process among two market participants or, for instance, the result of placing orders in an order book at an exchange. The triggering of one or multiple trades may then occur ad hoc or at prespecified points in time.

In continuous trading, a trade occurs whenever two orders match. Consequently, a mechanism must be in place that constantly checks for a situation in the order book that will allow this to happen. Every new order that reaches the order book is tested to determine whether such an order exists on the other side of the market.

Periodic call auction trading demands less effort than order book trading in a continuous market. This is because call auctions are inherently less complex. Call auctions are typically triggered whenever a certain, predetermined point in time has been reached (e.g., the opening of trading in the morning, the closing of trading in the evening, or at midday). Accordingly, auctions do not require a constant check of the market. Even triggers for volatility auctions as described previously do not come from the auction market itself; they arise in continuous trading, but only when a predetermined volatility condition has occurred.

In electronic trading systems, two types of triggers can be calculated, one depending on the order situation, and the other depending on time. Alternatives exist. In bilateral, negotiation markets, for instance, a trade occurs whenever one party to the negotiation accepts an offer that has been placed by a counterparty.

4.3.2 Determination of the Location of a Trade Occurs

Market models can be hybrids. In such situations, a combination of market models exists, and they may interact. Market models may be combined sequentially. For instance, in many markets trading starts with an auction at the opening that is followed by continuous trading, which is followed by an auction at midday, which is itself followed by continuous trading, and that then closes with a call auction in the evening. Depending on the order specification and market conditions, the order book that is eligible for the trade is determined.

4.3.3 Determination of the Counterparts of a Trade

Counterparty determination depends on how many parties interact with each other at the same time to find a trade. On one end of the spectrum (bilateral negotiations), counterparty determination can be relatively easy. However, searching for and selecting a party to start a negotiation might be costly. Negotiation starts when two parties enter a process of finding the details of a trade, including price, quantity, and post-trade modalities. That process ends successfully after the passage of some time, or it terminates without any result.

The market model is more complex when many parties interact simultaneously in the same place, and priority rules are imposed on all of them. If there are multiple buyers and sellers at the same time with orders in the market and their orders are all eligible to trade, priority rules are required to specify who gets to trade first. As such, the rules determine who trades with whom. The orders that are submitted to the market must carry certain requisite information. In most cases, price is the primary criterion used (the most aggressive orders trade first). Price priority (the primary rule) is typically followed by time priority (a secondary priority rule), and this requires that each order be time stamped when it enters the market. Of course, this in turn requires the mechanism of a clock that imposes a sequence on all orders coming in. Ideally that clock is a central mechanism positioned at the "gate" that all entering orders have to pass through.

The continuous trading and periodic call auction market models illustrate the difference between sequential, multiple-price, and bilateral matching type of trading on the one hand and simultaneous, single-price, and multilateral matching type on the other.

4.3.4 Determination of the Price of a Trade

The complexities and challenges of price determination are addressed in the previous section of this chapter. Notably, some market models exist that do not comprise price determination but which still lead to trades. Venues that follow such a model base trades on prices that are taken from other markets. The market model in these situations must define where to take a price from, and what specifically the reference price is; this is required for the production of either bilateral or multilateral trades.

Compared to a market model that is designed to produce prices based on the information that has been received and the rules that must be followed, building and operating a market model that comprises a reference pricing principle are clearly less complex.

4.3.5 Determination of the Quantity Traded

Price determination and quantity determination are two closely related functions that market models define. In those market models which include price determination, the quantity (i.e., volume of equities) that is offered (sellers) or sought for (buyers) depends on the price which buyers are willing to pay and sellers are demanding.

Both values—price and volume of transactions—in these cases are determined simultaneously. In automated order-driven models of trading those two functions are conducted algorithmically.

Market models which don't include price determination and use reference prices comprise a function for the matching of volumes to buy and to sell either periodically or continuously. Volumes in the first case are matched over a certain period of time and then executed with a quantity equal to the lower of the two quantities (buy quantity, sell quantity) available at that time. In case of a continuous matching a newly incoming order's volume is checked for execution against already "waiting" volume on the other side of the market or is going to wait until matching volume is available.

4.3.6 Determination of What Is Traded

Market models differ with respect to the standards they set for the assets to be traded. In highly standardized markets, all parameters that define an asset are preset and agreed upon. Traders accept and commit to these standards when they enter the marketplace. Security exchanges are of this type: their products are highly standardized and homogenous. A different situation exists when a market structure does not define, ex ante, the specifics of the assets to be traded. In these situations, market participants must agree on these qualities in the course of agreeing on a transaction, and the information exchanged must be specified accordingly.

4.3.7 Determination of Post-trade Modalities

Transactions that have been agreed to by market participants must be fulfilled, and the time and place for this to be done must be agreed upon. In exchange-type markets, that point is prespecified; typically, it is the national *Central Securities Depository* (cf. Chap. 6) where settlement takes place. Today, in most cases, that comprises the simultaneous, irrevocable, and final transfer of assets between the participants in the trade (typically money against securities). In cases where market structure does not specify that procedure, market participants must agree about how they want their obligations stated and fulfilled. That comprises finding a common place to which both parties can send their instructions for delivery and payment (i.e., the settlement of their trade).

The post-trade structures of exchanges standardly comprise the involvement of a central counterparty (cf. Chap. 5). For off-exchange transactions, regulators may demand the involvement of a CCP that offers clear standards to deal with consummated transactions.

4.3.8 Provision of Market Transparency Pre- and Post-trade

Market models specify the dissemination (or lack thereof) of information that reflects the trading intentions and orders of buyers and sellers. A market model transforms both the stream of incoming orders and the sequence of trades produced into information streams that reflect both orders and trades. The degree and timeliness of information available about pre-trading intentions are referred to as *pre-trade transparency*. The degree and timeliness of information available about trades that have occurred are referred to as *post-trade transparency*. Continuous trading in a so-called open limit order book fully provides both pre- and post-trade transparency. Information concerning volumes offered and requested at all price steps in the market is visible in real time to all market participants and to the public. The information about trades (i.e., price, the volume traded, and the exact time of the transaction) is published in real time as well. The term “lit-trading” is used for venues that, in this regard, are transparent. The term “dark-trading” is used by market participants who avoid (for various reasons) the publicity of lit-trading. Large traders, in particular, do not wish to have information about own trading intentions and completed transactions conveyed to other participants.

Chapter 5
Clearing

Thomas Laux

5.1 Introduction

Once a trade is executed, further actions are essential to finally settle it. At this point in the value chain—that is, post-execution—clearing is the next step. Clearing is the general term for the risk management and operational processing needed to ensure that the commitment made from the trade execution is concluded on the settlement side, a final step involving the exchange of shares for money and other critical procedures. In this final step, a central counterparty, or CCP, steps in between the original trading parties, becoming the new seller to the original buyer, and the new buyer to the original seller.

In this light, clearing is an indispensable part of a trade's life cycle, a piece of the plumbing which must operate smoothly to fulfill the trade execution. Clearing can be anything from a very light process to a heavy-handed process, depending on the maturity, breadth, and nature of a market. The former will lead directly to delivery instructions, the latter relying on the interactions among large financial market infrastructure.

Clearing has two critical features: The first is the transaction processing of the trade until the moment of settlement; the second is counterparty credit risk management between the original parties to the trade. Clearing can be completed bilaterally between the original counterparts to the trade, or it can be centralized in an entity on behalf of the trading parties. There is substantial efficiency to be gained from the participation of a larger group of participants. Most established regulated markets are centrally cleared and there is an increasing importance of centralized clearing to the OTC markets.

T. Laux (✉)
Eurex Clearing AG, Mergenthalerallee 61, Eschborn 65760, Germany
e-mail: thomas.laux@eurexclearing.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_5

Generally speaking, a financial market infrastructure that performs the first role—transaction processing—is called a *clearing house*; the second role—risk management—is performed by the *CCP*. Typically, however, these terms are used interchangeably, as most CCPs are clearing houses and vice versa.

The value in both roles is more pronounced with increasing market activity. For example, efficiency and automation are necessary for clearing if millions, or even thousands, of trades are concluded every day, across dozens of counterparties. A CCP provides substantial value in centralizing these functions. Furthermore, for contracts with less direct settlement, such as derivatives with longer maturities, the counterparty credit risk mitigation of a CCP becomes a critical feature. Once a CCP is established to facilitate all these features, it has certain other benefits that enable, for instance, multilateral netting and post-trade anonymity.

In this chapter, we focus on the clearing of trades through a CCP. The CCP model for cash equities has many features in common with other asset classes. However, in this chapter we consider central clearing in a broader context. The reason is not only because of the recent regulatory focus on systemic risk management. Also since the effects of clearing are amplified for trades such as futures and options with longer settlement periods. Moreover, this broad sweep is important given the relevance of the derivatives markets in price and information formation in relation to the underlying equities.

5.2 Background

5.2.1 History of Clearing

The term "clearing" is generally used for the process of matching claims. In a financial context, clearing first denoted both the sorting and settlement of claims in payment systems. Examples from history are the forerunners of central banks arranging the transfer of money across their member banks. Around the turn of the nineteenth century, the development of clearing houses and central counterparties was closely related to the development of organized commodity markets, especially in the USA and in Europe. In the modern context, clearing is often used for processing trades in financial instruments and products between CCP members. Most recently, risk mitigation by the CCP to the trading parties is the focus of this portion of the value chain, followed by the rise of electronic trading and clearing on the operational side.

Not surprisingly, this advanced technology now allows exchanges and traders to transact millions of trades a day across the globe. By contrast, in the 1960s, as the physical sorting of paper led to a bottleneck (often referred to as the paperwork crisis), the New York Stock Exchange held a limited trading session on Wednesdays to enable more time for clearing to be completed. Since then, the industry's clearing house, the Depositary Trust and Clearing Corporation, or

DTCC,[1] has evolved to deliver highly reliable settlement services, handling obligations on a near-time basis. Meanwhile, the growth in services by exchanges and clearing houses, as well as the internationalization of trading, has led to a consolidation of CCPs globally in the major markets.

5.2.2 History of Central Counterparties

Recent Developments—The crisis and clearing obligations, market structure changes.

The financial system was shaken by multiple shocks during the financial crisis that began in 2007. And this crisis put counterparty credit risk in the spotlight because of the systemic implications. Policy makers and regulators were soon at the forefront, seeking to shore up systemic risk mitigation, in particular by broadening and strengthening the use of CCPs. In doing so, entities whose operations were often overlooked moved to the top on the regulatory agenda.

Most importantly, G20[2] leaders agreed to mandate central clearing for certain types of trades. The CCPs today are not exactly a new phenomena. CCPs similar in concept to today's have been used in various futures markets since the beginning of the twentieth century. The OTC markets, growing rapidly since the 1980s, were generally cleared bilaterally. Still, the regulatory drive to strengthen CCPs will help offset a market structure weakness today—the lack of adequate risk mitigation. This important drive has fortified the standing of the CCP in the financial system.

While the major jurisdictions adopt legislation and regulations on clearing, rules have been simultaneously drafted to define more precisely what a CCP is and how it should be organized. On the global scene, *CPSS-IOSCO*[3] drafted the Principles for Financial Market Infrastructures. These principles are reflected and often enhanced in national laws. The primary regulation that influences clearing in the European Union/European Economic Area is the European Market Infrastructure Regulation (EMIR). In the USA it is the Wall Street Reform and Consumer Protection Act (Dodd-Frank).

New regulation elevates the importance of CCPs in the market today. We consider how their crisis management role today is often reflected in insolvency laws.

---

[1] http://www.dtcc.com/.

[2] The Group of Twenty (also known as the G-20 or G20) is an international forum for the governments and central bank governors from 20 major economies.

[3] The Committee on Payment Systems and Settlements, since then renamed to Committee on Payments and Market Infrastructures, and the International Organization of Securities Commissions are international bodies which bring together national and regional regulators and central banks to coordinate and outline joint policies and standards.

5.2.3 Global Context and Focus

The CCP landscape has evolved with the markets. Today, major financial and commodity markets are served by a CCP for at least a share of their trading activity. The preeminent CCPs are CME Clearing, DTCC, Eurex Clearing, LCH.Clearnet, ICE, OCC, and SGX. There are more local or specialized CCPs, such as the cash equity-only clearing houses in Europe, for example EuroCCP.

Exchange groups in Asia, with notable exceptions, tend to serve only their domestic markets. Depending on its market structure, a CCP can serve multiple or single markets or exchanges. The largest CCPs by volume, often with international capabilities, offer clearing for multiple asset classes. CCP clearing supports all major asset classes in both the cash and derivatives markets for equities, fixed income, commodities, foreign exchange, and credit. Many of the largest and most advanced CCPs are part of exchange groups, especially groups operating large derivatives markets. Derivatives markets have historically required more prudent counterparty credit risk management. For example, the potential rise and fall in value of long-dated contracts required "*performance bonds*" to be pledged to prevent trading parties from walking away from unprofitable agreements.

5.2.4 Definitions

Key institutions in the clearing process:

- *CCP—Central Counterparty*. Legally, the seller to every buyer and the buyer to every seller.
- *CH—Clearing House*. An organization which performs the sorting and filing of claims, often by establishing a net amount to be paid by every member or participant.
- *Trading Locations*—Trading venues from which the CCP obtains or CH takes up positions to manage. A variety of new execution platforms are now linked to CCPs, in addition to the traditional, strictly regulated exchanges. For OTC swaps, these include swap execution facilities (SEF) in the USA, organized trading facility (OTF) in Europe, or multilateral trading facility (MTF) for typical share trading in Europe. Away from an organized market, CCPs and clearing houses can also interpose themselves between trades concluded bilaterally.
- *(I)CSD—(International) Central Securities Depositary*. An organization that holds physical or dematerialized assets such as shares and bonds on behalf of its customers. These accounts are often the ultimate source and destination of trades.
- *Clearing Member*—A direct member of a CCP, often a bank or broker/dealer that may trade on its own account or provide clearing services for clients.
- *Client*—Not a direct clearing member itself but the ultimate trading participant that clears through clearing member.

5.3 The Clearing Process

5.3.1 *From Trading to Settlement*

The term "clearing" has acquired an official definition with recent regulatory overhauls in finance. EMIR, for instance, describes it is as "… *the process of establishing positions, including the calculation of net obligations, and ensuring that financial instruments, cash, or both, as available to secure the exposures arising from those positions.*" The US Commodity Futures Trading Commission (CFTC) defines clearing as "*A derivatives clearing organization (DCO) is a clearinghouse, clearing association, clearing corporation, or similar entity that enables each party to an agreement, contract, or transaction to substitute, through novation or otherwise, the credit of the DCO for the credit of the parties; arranges or provides, on a multilateral basis, for the settlement or netting of obligations; or otherwise provides clearing services or arrangements that mutualize or transfer credit risk among participants.*"

5.3.2 *Trade Entry into the CCP*

Once concluded, a trade then moves onto the clearing stage. A central counterparty agrees to clear trades either by "open offer" or "novation." In either case, the CCP becomes the seller to every buyer and the buyer to every seller, replacing the original trading parties as the counterpart to the trade. Open offer means that the CCP automatically steps into the trade. This is the generally accepted model for regulated markets that the CCP rules stipulate in the kinds of trades it will clear. The rules of this market—among them checks of price reasonability, maximum order quantities, as well as the clear definition of the instruments or products traded—enable a CCP to interpose itself without other checks of its own (Exhibit 5.1).

Novation is the validation by the CCP of a trade's characteristics before the CCP can accept the trade and step in. Novation typically is used when a contract or trade, an OTC trade for instance, is already agreed. At this stage, it is torn up, and

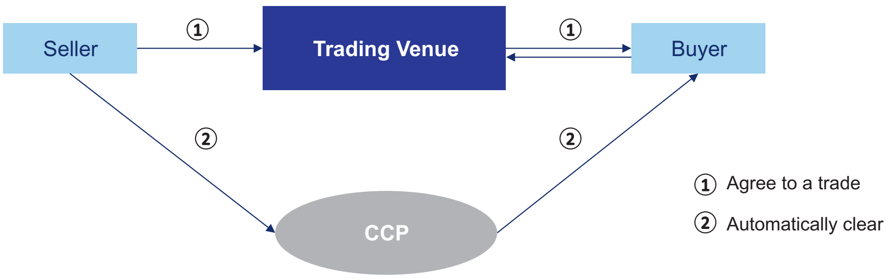


Exhibit 5.1 Open offer

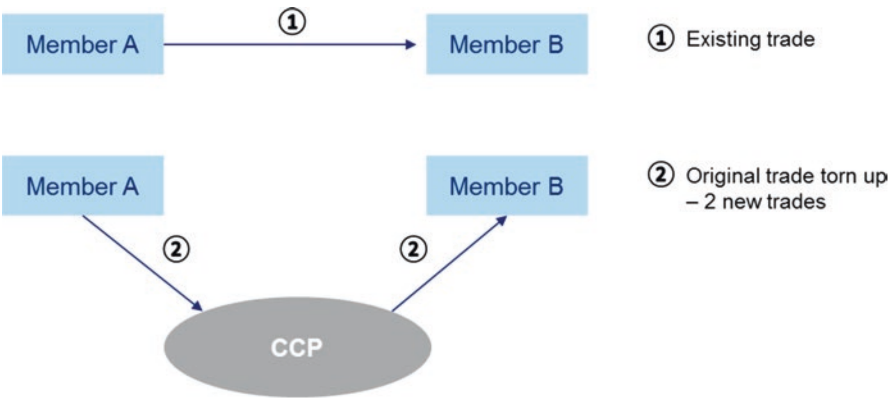


Exhibit 5.2 Novation

replaced by two new contracts or trades between the trading parties and CCP with identical contracts. If, say, an *interest rate swap* (IRS) is agreed between two counterparties OTC, the CCP will verify that the details of the trade are conformed to the product offering, before it is accepted by the CCP which clears them. These details include the currency of the trade, the interest rates, remaining maturity, and other specifications. Additionally, a CCP may have risk-based measures in the novation checks, verifying, for example, that the resulting margin requirement does not exceed a threshold.

Novation can in principle be performed manually based on term sheets. In practice, however, OTC CCPs conduct their novation electronically based on input from swap execution facilities/organized trading facilities and other affirmation/confirmation platforms. Examples include Bloomberg and Tradeweb.[4] A competitive market of vendors, fostered by the advent of OTC market regulation, are joining the traditional inter-broker-dealers in the trade matching space (Exhibit 5.2).

Two basic business models exist for clients: the principal-to-principal model and the agency model. The former means that—in addition to the legal trade[5] between the CCP and its clearing member—the client has an equivalent legal trade towards its clearing member, with the clearing member acting as principal from the client's perspective. In the latter model, the clearing member guarantees the trade for the client who has a legal trade with the CCP. The principal model is customary in Europe, the agency model being the US standard.[6]

A single point of concentration for trade processing to settlement is created with the CCP interposed as the counterparty in all transactions. This in turns leads to substantial efficiency since a net sum can be created per each member.

[4] http://www.tradeweb.com/

[5] Legal trade means that there is a contractual obligation between two parties.

[6] In the agency model, the end user of the trade legally holds a position at the CCP. The direct clearing member guarantees the trade and acts as agent of the client towards the clearing house, wherefrom the name.

In modern electronic markets, this is especially beneficial where hundreds or even thousands of traders across firms trade millions of instruments and contracts across the globe daily.

5.3.3 *Participants of the CCP*

A CCP must verify that both (a) the trade is in its product offering and (b) the parties are known that it is interposing between. For this purpose, a CCP maintains a list of accepted "members," or direct participants. These members maintain accounts with the CCP, and must meet the access criteria. Indirect participants (that is, clients of members of the CCPs) can also have their trades cleared by the CCP (Exhibit 5.3).

The CCP must retain records and any obligations of the cleared trades. The trade information must be processed and entered into the CCP. The CCP requires its members to keep the following accounts:

- Cash accounts: The cash accounts are where the members hold money required for the settlement of trades. These must be fully funded, and accessible to the CCP if, for instance, they are maintained at third-party service providers such as banks, custodians, or central banks.
- Position accounts, including gross accounts and market making accounts. These track all the trades a member or its clients have with the CCP.

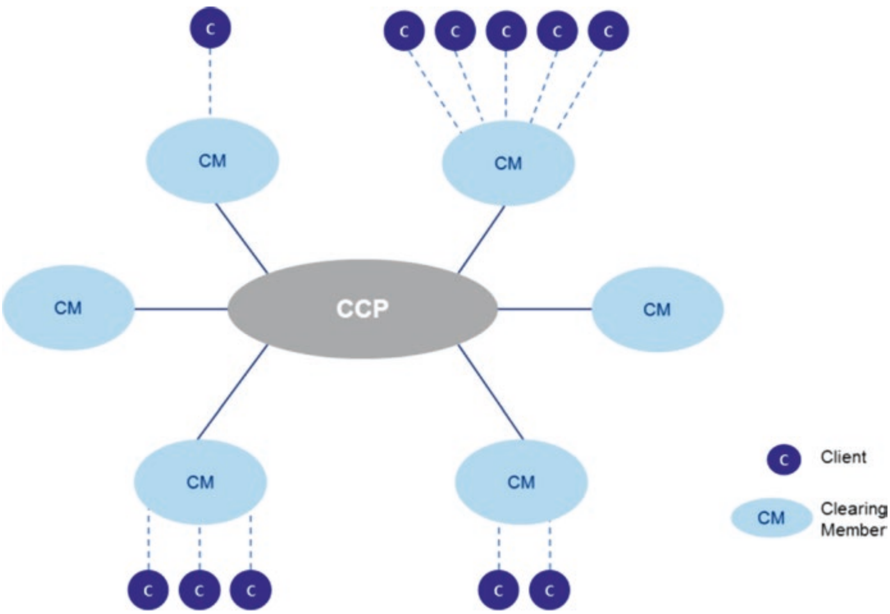


Exhibit 5.3 The clients, Clearing members, CCP

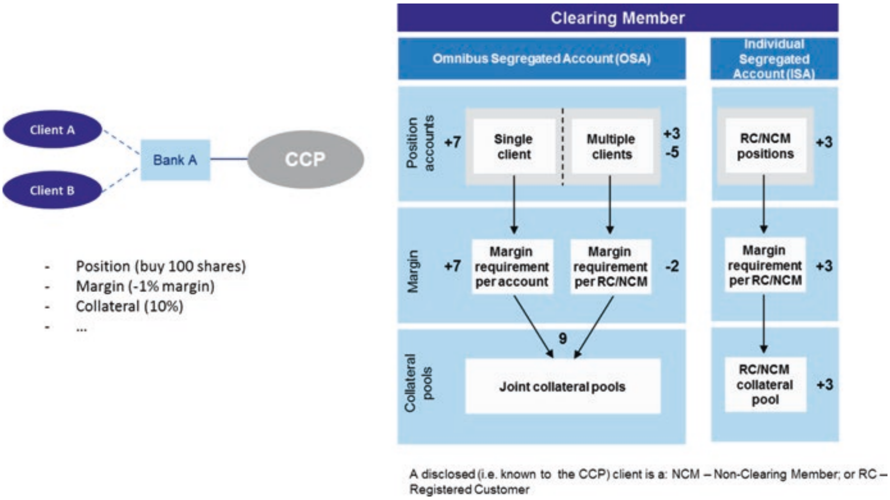


Exhibit 5.4 Account structures covering different segregation models

- Margin collateral accounts: These may hold collateral in the form of bonds, typically held as a pledge towards the CCP in case of member insolvency. The margin requirement set by the CCP is the amount that the account must hold to avoid a margin call.
- Settlement accounts: Accounts held at an (I)CSD by the clearing members and are used for the physical settlement of, say, shares.

Aside from any regulation for trading and client business, these accounts may require clearing members to be regulated as banks or other credit institutions. The accounts are typically held at either a credit institution accepted by the CCP, a central bank, or an (I)CSD. The (I)CSD may also be the organization that performs the ultimate delivery-versus-payment instructions for settlement. If a CCP facilitates multicurrency trading in instruments or products, this must be supported by the appropriate accounts and settlement infrastructure through central bank money, or at commercial banks.

The account structure has a second layer since client trades must be separated from the proprietary trades of the clearing members. CCPs, traditionally, maintained only segregated position accounts for their clients because the margins and settlements were both provided by their clearing member. Client accounts are segregated to mitigate the credit risk of their clearing member (Exhibit 5.4).

These generic account structures may vary for specific markets. In centrally cleared repo markets, for instance, it is uncommon to have client business. As a result, all trading participants are typically direct members. On the other hand, Dodd-Frank and EMIR stipulate different segregation models: in the USA “LSOC”[7] is the only choice; in the EU/EEA clients can choose between omnibus segregation models or individual ones.

[7] LSOC = Legal separated, operationally comingled, i.e., a form of “tagged” omnibus segregation.

5.3.4 Netting and Offsetting

Netting of trades is the process of reducing outstanding buy and sell orders on both sides of pending trades into a single net figure. Offsetting is, strictly speaking, the process of creating a "single" trade out of gross transactions. The term "netting," often used for both, is, in effect, the sum total of all individual long and short positions in the same instrument or contract. Trade netting helps exchanges and clearing corporations to pool all the pending trades and settle them simultaneously (Exhibit 5.5).

There is, however, a clear legal difference. Netting does not replace individual transactions, but instead treats them jointly; offsetting cancels the opposite position in the same trades. Netting in this sense can be done bilaterally for any two parties with more than one outstanding obligation, assuming that they have the same delivery obligations. An offsetting trade simply takes an opposite position in the same contract.

Notably, the ultimate delivery obligation must coincide, even though trades at different prices can still be netted. For instance, if during a trade day two parties trade a share across each other multiple times, the final obligation is the offset and net amount. The prices at which the trades were executed must still be monitored and recorded (Exhibit 5.6).

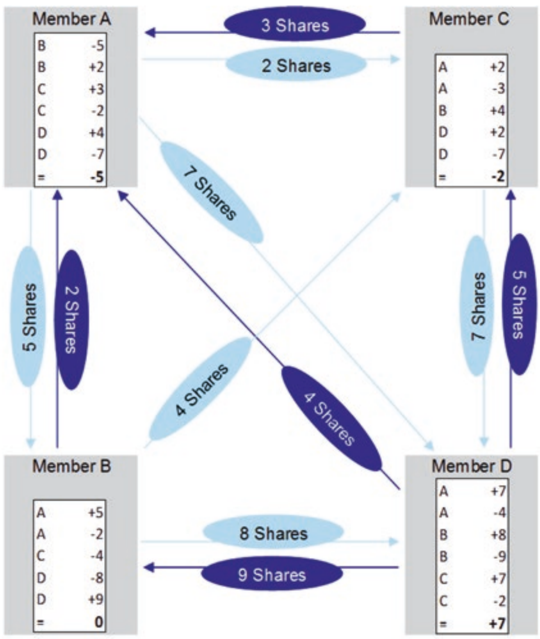
The concentration of clearing into a central organization produces a superior type of netting: multilateral netting. In this process, the CCP establishes a single outstanding obligation for each of its members individually, regardless of who they have traded with (Exhibit 5.7).

Both forms of netting minimize obligations, and hence both exposure and operational requirements. One reason why CCP risk management and operational processes can be so effective: the CCP has, per member, the minimum number of possible instructions to process. Each CCP tends to have a dominant share in a specific market partly because dividing trades in one asset across multiple CCPs decreases netting efficiency even more since multilateral netting cannot be maximized for the best performance. Generally speaking, netting efficiency of a CCP expands with more trading parties and daily volume in specific instruments and contracts. In the major markets, these can approach netting efficiency figures of over 90 %.

Even in un-cleared OTC markets, "compression"[8]—that is, the reduction of identical but opposite trades—has assumed more importance. Compression is a function similar to multilateral netting. Compression cycles are arranged by organizations, such as TriOptima,[9] and across a wide range of members. Not surprisingly, these are more cumbersome than the multilateral netting for fungible trades typical

[8] Compression is basically the replacement of (mostly) offsetting trades, and possible a large group of them, with a lower number of trades. In some cases, lower number is taken to mean smaller notional.

[9] http://www.trioptima.com/.

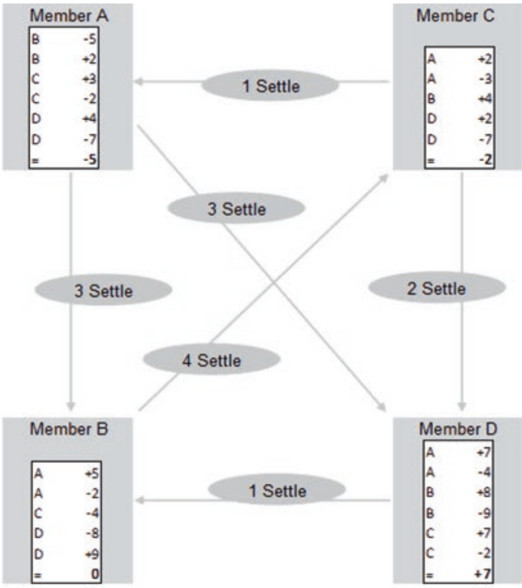


Trade	Impact on A	Result B	Result C	Result D	Settlement Y / N
A sell to B	- 5	+ 5			Yes
B sell to A	+ 2	- 2			Yes
A sell to C	- 2		+ 2		Yes
C sell to A	+ 3		- 3		Yes
B sell to D		- 8		+ 8	Yes
D sell to B		+ 9		- 9	Yes
C sell to D			- 7	+ 7	Yes
D sell to C			+ 5	- 5	Yes
A sell to D	- 7			+ 7	Yes
D sell to A	+ 4			- 4	Yes
B sell to C		- 4	+ 4		Yes

Exhibit 5.5 Buy and sell orders without netting

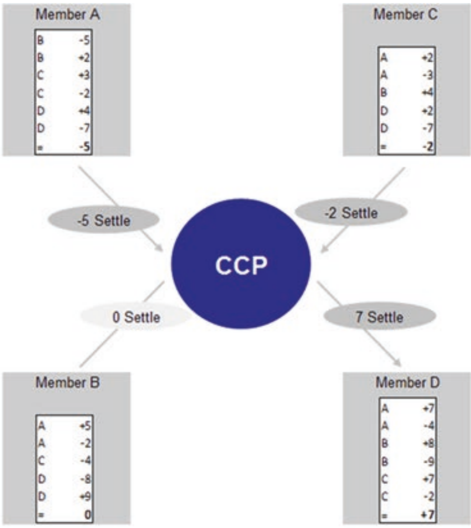
in the established CCP markets. Often, imperfection in the compression algorithm is tolerated. As OTC contracts increasingly come into the centrally cleared space, CCPs are adopting the compression services for their members.

Netting of pending obligations can be done either at the end of the day or intraday on a real-time basis. Market maker accounts are usually held net. In this case, opposite trades directly offset each other once concluded. In contrast, compression cycles for OTC trades may be run once or twice a year.



Trade	Impact on A	Result B	Result C	Result D	Settlement Y / N
A sell to B	- 5	+ 5			No
B sell to A	+ 2	- 2			No
A deliver to B	- 3	+ 3			Yes
A sell to C	- 2		+ 2		No
C sell to A	+ 3		- 3		No
C deliver to A	+ 1		- 1		Yes
B sell to D		- 8		+ 8	No
D sell to B		+ 9		- 9	No
D deliver to B		+ 1		- 1	Yes
C sell to D			- 7	+ 7	No
D sell to C			+ 5	- 5	No
C deliver to D			- 2	+ 2	Yes
A sell to D	- 7			+ 7	No
D sell to A	+ 4			- 4	No
A deliver to D	- 3			+ 3	Yes
B sell to C		- 4	+ 4		No
B deliver to C		- 4	+ 4		Yes

Exhibit 5.6 Bilateral netting



Trade	Impact on A	Result B	Result C	Result D	Result CCP	Settlement Y / N
A sell to CCP	- 5				+ 5	No
B sell to CCP		- 2			+ 2	No
CCP sell to B		+ 5			- 5	No
CCP sell to A	+ 2				- 2	No
A sell to CCP	- 2				+ 2	No
C sell to CCP			- 3		+ 3	No
CCP sell to C			+ 2		- 2	No
CCP sell to A	+ 3				- 3	No
B sell to CCP		- 8			+ 8	No
D sell to CCP				- 9	+ 9	No
CCP sell to D				+ 8	- 8	No
CCP sell to B		+ 9			- 9	No
C sell to CCP			- 7	+ 7		No
D sell to CCP			+ 5	- 5		No
CCP sell to D			- 7	+ 7		No
CCP sell to C			+ 5	- 5		No
A sell to CCP	- 7				+ 7	No
D sell to CCP				- 4	+ 4	No
CCP sell to D				+ 7	- 7	No
CCP sell to A	+ 4				- 4	No
B sell to CCP		- 4			+ 4	No
CCP sell to C					- 4	No
Net A to CCP	- 5				+ 5	Yes
Net B to CCP		0*			0*	0*
Net C to CCP					+ 2	Yes
Net D to CCP					- 7	Yes

* Zero shares have to settle if there is a balance in the cash side of the transaction.

Exhibit 5.7 Multilateral netting

5.3.5 *Transaction and Trade Management*

Trades and obligations require updating, until they are passed into final settlement. Trade management can include adjustments or corporate events, in other words, changes to the trade. For cash equities, the typical settlement period in Europe is T + 2. Jurisdictions formerly T + 3 are moving to a shorter period as part of international harmonization. At the extreme end of the spectrum for derivatives, CCPs accept swaps that have maturities reaching 50 years. Still, the most common maturities traded are less than a year for listed derivatives, and around 10 years for OTC. Aside from the CCP being up to date on the trade, information on their exposures must be disseminated to members and participants of the CCP. All CCPs must have clear and definitive reporting, mostly by electronic dissemination, to their members on exposures and trade information. This devotion to record keeping can be extremely useful for regulators, particularly in times of financial stress when the CCP and its regulators can rapidly pull up figures on the open exposure of members and their obligations. The transaction and trade management of a CCP is completed once the obligations between the original trading parties have been settled. To this end, a critical aspect in deciding when the trade comes to and end at the CCP is settlement finality. Settlement finality is also critical for establishing legal clarity on the obligations from the trade execution. Without these finality rules, an insolvency administrator, for example, in the event of a default of a trade counterparty, could review and "cherry-pick" from outstanding obligations, claiming some to be disadvantageous or unfair. Strictly speaking, CCPs hold margin only on behalf of open obligations, so final settlement means that the members' margin collateral for the trades can be withdrawn.

Last Word: The clearing process as described here can be complicated and intricate, undoubtedly so on complex products. However, centralization and automation of the processing is a substantial cost saver if the markets are sufficiently active to justify the outlay. In these cases, CCPs benefit from the scale of operations, enabling trading to occur quickly and efficiently across a diverse set of firms.

5.4 Risk Management

In addition to processing the trades, CCPs assume counterparty credit risk management. Members are therefore guaranteed that concluded trades are settled as agreed. In doing so, a CCP must have strict risk management of both the members and their exposure, as well as resources in the form of collateral. If a member defaults on its obligations, a CCP must mitigate the situation and guarantee that the claims of counterparties are satisfied. A CCP fulfils its most basic obligations by requiring each member to contribute funds—known as *margin collateral*—for covering a member who fails. All the members contribute to a mutual fund for the CCP for the successful conclusion of open trades in the event of realized losses exceeding the margins of a defaulting member. In sum, this multilayered approach, the mainstay of the risk management framework, establishes incentives for strong risk management standards for the CCP and its members.

CCPs tend to be very conservative risk managers because of their incentive and disincentive structure. If a member defaults, for instance, the members can lose for example on their portfolio from the rebalancing of the CCP. However, they do not trade on their own account. And they have no upside from the trades. Surplus collateral is returned to the insolvency estate. Additionally, as the running costs are incurred by the members with the riskiest positions, the mutualization aspect of the CCP creates incentives for its members, especially the large ones, to favor strong risk management.

5.4.1 Novation and Pre-trade Risk Tools

In the clearing process, a CCP accepts trades either on an open offer basis, or novates trades between its members. Open offer is employed for listed markets, and the contract details are standardized and known to the CCP. Novation is performed for OTC contracts and the CCP conducting checks, including a risk check, before it accepts a trade and steps in as the new counterparty. These checks validate that the more bespoke trade matches the conditions for trades that the CCP accepts to clear as part of its service. The risk check can include a collateral level confirmation so that the incoming trade will be sufficiently backed without triggering a margin call. In recent years, clarity and immediacy, understandably subject to appropriate checks, have been tightened for the highest degree of certainty by a CCP in accepting a trade.

5.4.2 Counterparty Credit Risk: Access Criteria

A CCP, for either novation or an open offer, accepts trades only from its members, or clients of members as arranged legally through a member. The CCP members must follow criteria that satisfy the CCP's risk management standards, witch typically include:

- Basic legal requirements such as a regulatory status permitting the type of business performed
- Risk-based requirements including capital charges, minimum contributions to the *waterfall*,[10] and rating-based assessments
- Operational and technical requirements to make certain that members can perform the clearing processes; among the requirements, usually, are staff skills, infrastructure for trading and clearing technology, and ability to time-sensitively handle margin calls and settlements

[10] The CCP's "waterfall" is the sequence in which financial resources provided by the members, CCP, and possibly other stakeholders are used to cover losses from a member default.

- Other criteria, most often related to the default management processes of the CCP, such as mandatory participation in auctions

The credit quality of its members is a key consideration since the CCP is primarily a counterparty credit risk management entity. A credit risk assessment of the new member is performed before the member is joining a CCP. Once accepted as a member, the CCP performs regular evaluations of a member—its share prices, credit default, or bond spreads—and actively follows market news and information and conduct audits for both credit and operational processes on a member. The purpose is to keep abreast of any developments. In this way, a CCP can be prepared either to enact limitations to the business or risk that a troubled member can undertake at the CCP or to ascertain whether a credit event has happened or appears likely.

The types of events which cause the CCP to undertake a formal review or actions include the "trigger events" defined in their rules and regulations. Not surprisingly, other members have a keen interest in CCPs keeping a vigilant eye on the entire membership's credit risk profiles and their evolution.

As we shall see, the CCP depends on its membership to smoothly risk manage a default. In an active market, it is beneficial to have a broad and diverse membership to prevent a single, firm-specific crisis from affecting all the members simultaneously. However, in doing so, the CCP membership can have heterogeneous credit quality, ostensibly to the detriment of the safest members. These balances must be reflected in the risk framework of the CCP, delineating how much mutualization is permitted.

Historically, credit ratings are not a reliable guide to the likelihood of default, since these are often rare events. Put differently, if a well-operated CCP can substantially mitigate counterparty credit risk for its members, then the members benefit from being able to trade against each other—or a client of a member—without having to conduct due diligence, establish credit limits, and arrange bilateral collateralization requirements. For CCPs with international reach, the added benefit of cross-border business—and only having to consider the legal location of the CCP—can greatly simplify the process as the legal complexity is reduced. In such cases, the CCP must, of course, ensure that its rules and regulations apply in the jurisdictions of its members. But this centralized work is cheaper for the overall market, than it is done multiple times in a decentralized setup.

The large CCPs typically have 50–200 direct members, each with various types of license types. A considerable portion of the global financial markets is accounted for by these large CCPs and their clients. The benefits resulting from such a broad participating in the CCP "club" are otherwise offset by the higher probability of default, for at least some members. To this end, CCPs with varied membership tend to respond with higher collateralization and lower mutualization requirements per member compared with other CCPs. This lower mutualization is the lower proportion of the default fund contribution per member relative to their margin requirements.

The clearing members of a CCP are responsible for managing the default of their clients, regardless of whether or not they are agent accounts or exchange participants. In this sense, the clearing members protect the CCP and the other members from their clients, who may be very different entities to the typical member. For instance, they may be corporate hedgers.

5.4.3 Default Management Process

Despite the admission criteria, members of CCPs may still fail to perform their obligations. This may be due to short-term operational reasons or to the bankruptcy of the member. A CCP must ascertain the cause of a member's failure and if it can be satisfactorily mitigated, so that the CCP is not required to step in to perform on its guarantee of the trade towards non-defaulting members. To distinguish between such cases, CCPs have a list of trigger events in their rules that outline how to respond to certain situations.

Trouble at a CCP might also spell general problems for a member. In the case of bankruptcy, as CCPs declare member defaults, care must be taken so that a healthy member is not erroneously called out on defaulting. Indeed, this error could result in considerable secondary effects across the markets. CCPs often check regulatory communications before declaring a default. New CCP legislation requires CCPs to inform a supervisory or prudential authority[11] if a member defaults.

If, however, a member is actually in default, for example, because of insolvency, then the CCP must act rapidly to ensure that smooth CCP operations continue and that the risks associated with the default are managed appropriately.

CCPs must prudently manage the risks, serving the market in an operational sense. The main risk of a CCP, however, is a member default. If a member defaults, trades originally between the defaulting member and the CCP are terminated. The CCP subsequently becomes unbalanced (Exhibit 5.8).

In this case, the CCP assumes the responsibility of settling the trades against the non-defaulting parties on the other side of the trade. The CCP is also exposed to the resulting market risk. A CCP inherently has substantial market risk, contingent on a credit event occurring for a member. The usual response is that CCPs assume a default probability of 100 % for all of their members, regardless of credit rating or their own internal statement of credit quality. To cover for this risk, the CCPs charge margins based on the market risk of the member's portfolio to a high degree of confidence. The contingent market risk that would arise from a default is covered to a high degree of certainty on a constant basis, as if default was to happen tomorrow.

The default management process is, in effect, the art of reestablishing a balanced book by the CCP. The CCP's default management process must protect the integrity of the CCP and use the minimum possible funds from its waterfall. In doing so, the CCP's rebalancing shields the non-defaulting members of its markets from adverse impact arising from the default, enabling orderly trading to continue across the

[11] A prudential authority is responsible for oversight of capital markets and their actors in a macro-sense. Often, a prudential authority defines its responsibilities in terms of "systemic risk," that is to say, risk which is realized through the interaction between various institutions and their behavior, rather than the (micro) regulation of a particle entity. Following on from the financial crisis in 2008, various systemic risk boards, such as the FSOC or ESRB, have been created, and other regulators, such as the Bank of England, have developed a function to address this feature of markets.

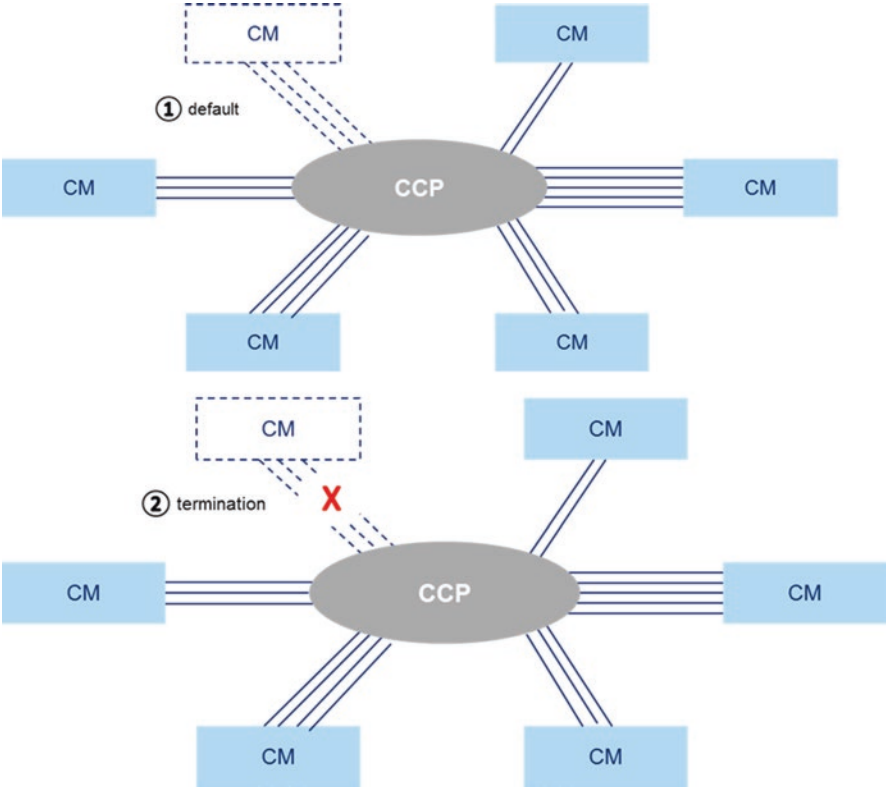


Exhibit 5.8 Termination of trades

markets. As CCPs must have the ability to ensure that they can rebalance towards their non-defaulting members, their rule book makes the non-defaulting members act as mandatory purchasers or sellers.

The default management process depends on the type of positions the CCP must rebalance. For regulated markets, the CCP will typically conduct trades on the primary trading venues. CCPs, especially so for listed derivatives, are often part of exchange groups, the primary market being the mother company or a part of the same group as the CCP. For OTC contracts, CCP liquidation procedures must follow market conventions. Often, they have a default management group of external traders to assist them. Such external traders are usually seconded[12] from the largest members of the CCP.

Default management can include auctions in addition to trading directly to rebalance the CCP's books. Members (and sometimes their clients) in an auction bid for the defaulter's portfolio. Auctions are either voluntary or mandatory. In the latter,

[12] Secondment means the temporary movement or placement of an individual into another role or position.

members actively trading in the positions to be auctioned are often selected from the CCP's members. If an auction is unsuccessful, some CCPs will invoke rules to ultimately guarantee a matched book. There are rules such as "invoicing back" or "allocation."[13]

Auction formats must be designed so that they can enable good price discovery. If multiple positions are included in a portfolio auction, care must be taken in selecting the positions to be jointly liquidated. If the auction portfolio is too large, members may struggle to provide a price, especially if they are not active in some of the instruments or contracts. On the other hand, a single auction enables the CCP to create a market and liquidity for positions in aggregate. From an operational perspective this is an efficient approach. Prior to auctioning a portfolio, a CCP can hedge the defaulter's positions, therefore reducing the risk of losses for the CCP as well as making the auction portfolio easier for the bidders to price. Hedging is generally based on reducing the portfolio sensitivities of the portfolio, focusing on the most important ones, and allowing for basis risk to remain until a later stage. The sensitivities that need to be hedge depend on the defaulter's portfolio.

Directly trading out of positions, hedging, and auctioning are time consuming. By definition, the more liquid a market, the faster a CCP can rebalance positions without undue price impact. As market participants will charge a liquidity premium for larger trades, often to reflect the resulting inventory risk this creates for them, a CCP rebalancing itself for larger positions will typically result in an inferior price as the CCP pays for immediacy. The longer the liquidation of the positions, the greater the possible variance in the portfolio's value. While the CCP could make profits from the portfolio, the plan is always to reduce risk. The reason: the purpose of the CCP default management is to protect the non-defaulting members. Therefore, minimizing the possibility of losses leading to mutualization[14] takes precedence. Of course, the CCP's own capital is also at risk (Exhibit 5.9).

To that end, CCPs must carefully evaluate the expected time it can take to exit positions, charging members collateral carefully calculated to cover possible losses. Once the basic parameters of an expected default scenario are established, the margin model can be determined.

When the CCP margins the exposure of a member, a choice must be made on the aggregation, ranging from the entire portfolio down to individual net trades. This balance of portfolio effects is undertaken in practice along asset class lines, or smaller. Any margin offsets which are granted must be realizable in the default management process.

---

[13] Invoicing back means that the original counterparty to the trade is requested to take up the position directly, essentially removing the CCP from any further obligations. Allocation often denotes a similar process, although trades are not returned to the original counterpart, but across non-defaulting members of the CCP. Allocation is often performed with the members taking over obligations to rebalance the CCP in proportion to their activity.

[14] Mutualization is the use of common funds, often called the default fund. Using funds from non-defaulting members is rare and considered a grave event.

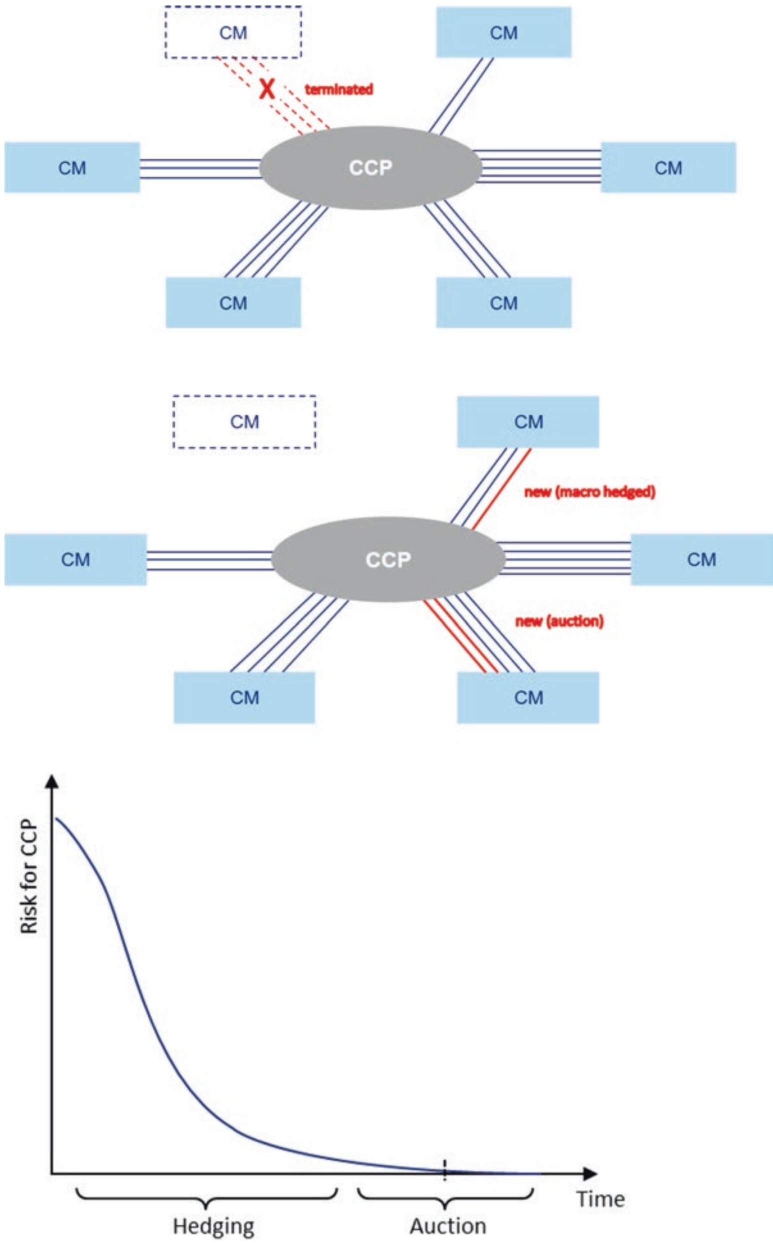


Exhibit 5.9 Default management process

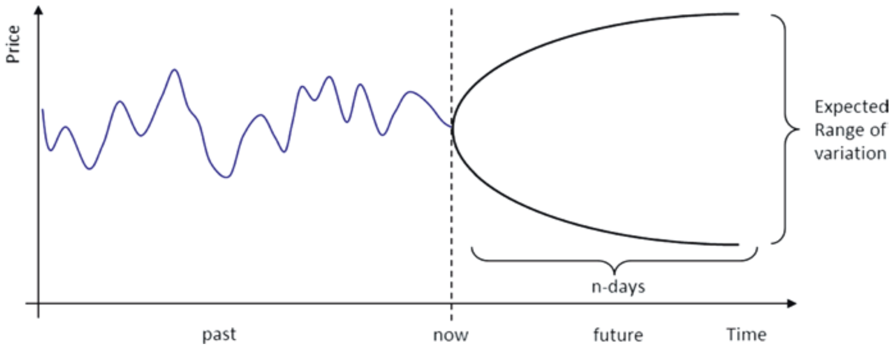


Exhibit 5.10 Holding period and price moves

5.4.4 Basics Margin

Margin is the primary risk mitigation tool of the CCP. It is collateral, in cash, securities, and sometimes commodities, collected by the CCP against the exposure of a member. Margin is composed of two components: mark to market and forward looking. The former captures the difference in price between the price at which the trade was executed, and its current market valuation. The latter is the CCP's expected loss—typically at a confidence level in excess[15] of 99 % for OTC derivatives—if it needs to liquidate the position. This confidence interval is computed based on the expected holding period, generally ranging from 2 to 5 days. In effect, this means that the CCP is fully covered for market moves with at least a 99 % confidence level over the holding period.

The holding period itself may vary by product, or the concentration add-ons[16] can be thought of as compensating for immediacy within the shorter time period. Generally, anything within 5–20 % of daily volume can be liquidated without undue endogenous effect on market prices and market liquidity. If the same or similar products are cleared at different CCPs, calculating the volumes and liquidity becomes more difficult. Listed derivatives exchanges report daily volumes on their website. Since these are often in the public domain, it is far easier to determine liquidity add-ons than for OTC contracts. For the latter, CCPs supplement their own volume figures from their members or vendor data (Exhibit 5.10).

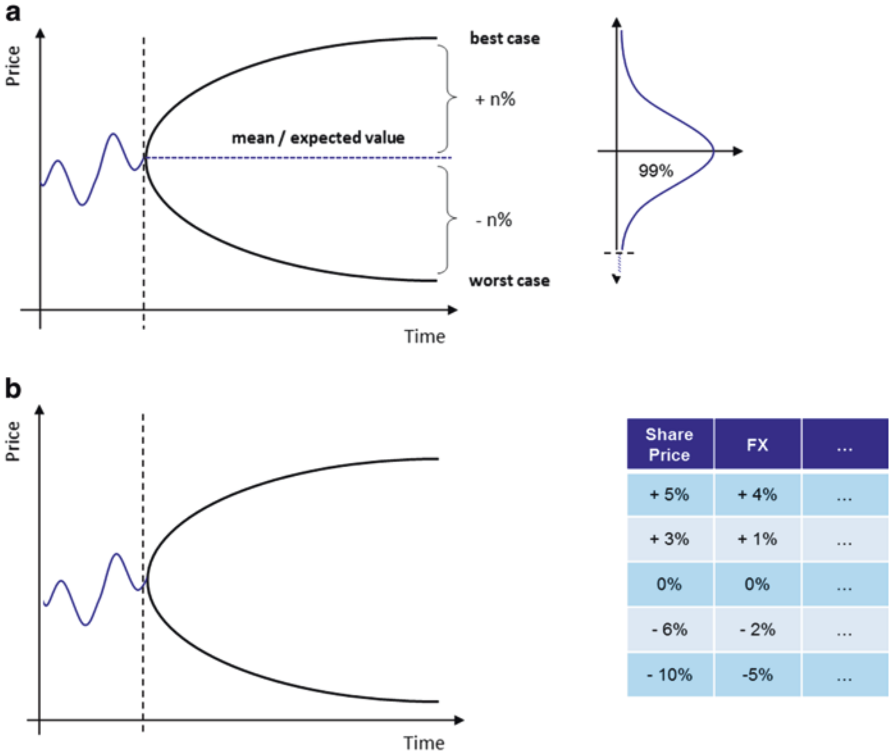
---

[15] Given the preference most CCPs and clearing members have for a "defaulter pays" model, most CCPs have higher confidence levels; for instance LCH.Clearnet uses a 99.7 % confidence level, and most CCP's back-testing reveals de facto higher confidence levels on a regular basis. The European Market Infrastructure Regulation requires that CCPs maintain at least 99.5 % confidence level for OTC derivatives such as interest rate swaps.

[16] A concentration add-on is further margin charged to account for the discount that trading out of a large position is expected to cause. For example, as it is harder to sell 100,000 shares than 10,000, a CCP should charge a member more per share for the first position. This key feature is important to keep orderly liquidations and mitigate losses beyond the defaulter's collateral.

Margin computation is a market risk methodology suitable for quantitative measurement if the products are liquid enough to have reliable price series. This can either be a simple value-at-risk approach or a scenario-based matrix model such as SPAN or RBM. The use of Monte Carlo methods, with the notable exception of the OCC's STANS, is rare, since members have a transparent and predictable method for how much they are charged. For these models, the CCP can compute margins either on a trade-by-trade basis, or on a portfolio level, according to its default management procedures (Exhibit 5.11).

A risk model requires data on even more extreme moves when it targets higher confidence levels. For instance, a 99 % confidence level covers all but the greatest loss of every 100 trading days. A 99.9 % confidence level covers all but the greatest loss of every 1000 trading days. So the quality of input data becomes an issue since the margin for these rare tail events is very sensitive to inputs. Quantitative modelling techniques are best practices to ensure robust margin figures for high confidence levels. This is typically achieved by making conservative assumptions about the distribution of tail losses. Back-testing is a natural check. But CCPs must determine sufficient margin requirements for worst-case, black swan scenarios since



Share Price	FX	...
+ 5%	+ 4%	...
+ 3%	+ 1%	...
0%	0%	...
- 6%	- 2%	...
- 10%	-5%	...

Exhibit 5.11 Margin range in a scenario-based matrix

these are often the cause and effect (even both) of the only scenario they must cover: a member default. Beyond this, a purely historical model makes the assumption that the future resembles the past, even when distance stresses in the past are maintained in the data set. Stress testing beyond the margin model is the main approach to this.

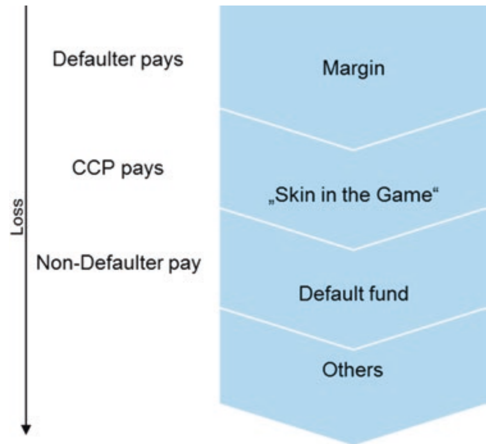
5.4.5 Lines of Defense

A CCP maintains other resources to cover potential losses from a default in addition to the margin collected from each member. These are collectively called the defaulter waterfall, or the *lines of defense*. The purpose of further resources is twofold: added safety to the CCP and aligning the incentives of other participants and the CCP. The cumulative effect: trades cleared by CCPs are guaranteed to a much higher value than only the 99 % confidence interval that a member's margins are set for. In fact, the credit quality of the CCP is substantially higher at a lower cost because the tail risk is covered by joint funds across all the members.

The lines of defense are calibrated to extreme but plausible scenarios. As a standard, they must cover the default of at least two clearing members and their clients. The two clearing members selected have the worst possible losses beyond their own margins in the stress scenarios. Covering two may not be the worst scenario. It is possible for more members to default in a short period of time. In this case, the CCPs may be covered, especially if the defaulting members' positions are offsetting to any degree (Exhibit 5.12).

The lines of defense can be calibrated to the desired confidence level, often at 99.9 % or above. However, this becomes more of an art than a science as we move further into the tail.

Exhibit 5.12 Generic default waterfall



Once the required tail-risk coverage is ascertained based on the input parameters, this must be collected in a joint fund and allocated to the members. This is usually conducted on the basis of a simple weighting according to the initial margin of each member, reflecting their size in the market.

Not surprisingly, the larger the membership and number of fixed scenarios and predetermined defaults the fund is intended to cover, the lower the charge per each member. At major CCPs, the default fund charge per member is in the range of 5–12 % of the initial margin, in addition to a minimum requirement.

The CCP also contributes to the lines of defense, a familiar business idea colloquially known as "skin in the game." This adds another incentive for the CCP to manage a default within the margins of the defaulter. This is because any losses exceeding it affect the CCP first.

In the most dramatic cases, a CCP often has other tools to either reduce its liabilities or raise more assets to keep the clearing operations, if the waterfall is insufficient. Typically, these are ways to increase funds available to the CCP from its members and third parties and its own resources. A common approach is further assessments of default fund contributions called from its members.

Oftentimes, the waterfall can only be used to cover losses from a default. The other possible losses by the CCP from, say, litigation or operational risk, are covered by the CCP's capital.

Last Word: As we have seen before, CCPs are very strict on risk management. They ensure that any accumulated profits and losses are settled or covered on at least a daily basis. They also charge members for their market risk on a very high level of confidence, assuming a 100 % probability of default. This transformation of credit risk to market risk enables a more accurate risk management, that is, provided that the products are sufficiently liquid for quantitative models.

5.4.6 Margin Calls

When market prices move, the CCP's participants' position value changes. For those participants whose positions have lost value, this unrealized loss results in variation margin shortfalls, where the CCP essentially offers credit to members. CCPs have implemented margin call processes to handle these credit scenarios as well as to ensure that it does not grow beyond predefined thresholds that form a sort of risk appetite of the CCP. A CCP must have technology suitable for computing position values rapidly and to input the relevant data to determine profit and loss intraday. This becomes more difficult in markets that are more fragmented or opaque than other markets. Here are the two extremes on this scale: trades on a regulated market, the CCP as the primary clearing house; and an OTC market with multiple CCPs. No matter what the source of the input data for the margin calls, robust procedures ensure that data errors do not cascade into an unnecessary margin call.

Once the profit and loss are calculated, a CCP and its members must have the operational capability to perform quick margin calls (often within an hour) of when losses exceed thresholds defined by the risk appetite parameters. These may be a minimum absolute amount, as well as a relative limit to either the member's existing exposure or the mutual parts of the waterfall.

The latest time frame CCPs settle the cumulative changes in position value is daily, in the so-called overnight margin call. Rigorous mark to market has the benefit of preventing losses from accumulating. And in the waterfall construct, large credit by the CCP is inappropriate. Nonetheless, this means that trades cleared by CCPs can require substantial liquidity from its members and their clients in times of stress.

Historically, initial margin at a CCP had been sufficient for member defaults. In extreme cases, such as Black Monday in 1987, the mark to market was problematic. In Chicago, certain members were late in delivering variation margin, leading to a CCP payout delay against the other sides of the trade. In contrast, the expected failures in variation margin payments following this same crash were the trigger for closing the market and an emergency recovery plan in Hong Kong.

5.4.7 Liquidity Management

Once trades are concluded, a CCP's members must have sufficient liquidity for the settlement of their positions, timely payment of variation margin to the CCP, and initial margin payments. In the conduct of its business, a CCP will need liquidity if a member defaults. Additionally, to ensure smooth processing and minimizing the impact of technical defaults by members, a CCP can also employ its liquidity lines outside of member bankruptcies. In both cases, the CCP does not need liquidity for the initial margin because this must be paid by the members.

These potential liquidity needs in the event of member insolvency are of paramount importance. The liquidity lines must be scaled to fit multiple member defaults, such as "cover 2." Liquidity lines from commercial banks and central banks and CCP's own funds are sources for liquidity. A CCP often has reverse repo arrangements, securities collateral being exchanged. Commercial credit lines might be unsecured; central bank money is often obtained only against collateral.

5.4.8 Collateral

Members and their clients must provide the necessary collateral once prudent margin and clearing fund requirements are determined by the CCP's risk models and framework. This collateral can consist of cash, securities, other instruments, and commodities of value. The CCP must ensure that if noncash collateral is deposited, it can be rapidly converted into cash and without loss in case of default. To this end, CCPs apply haircuts and concentration limits to the noncash collateral they accept. If cash is accepted in currencies different to the exposure of the risk of the portfolio, a CCP applies an exchange rate haircut to cover for the exchange rate risk. At the same time, it maintains the ability to trade one currency for another corresponding to the transfer amounts.

The legal nature of the collateral to CCPs depends on jurisdiction and the type of client. But the most common configuration for the prominent CCPs is that cash is title transfer[17] to the CCP, whereas noncash collateral is a pledge. Unlike default fund collateral, the margin collateral of non-defaulters is not used to cover losses of others.

What about collateral management? Cash collateral is often invested, the return divided between the CCP and the members who deposit cash. This is a standard practice because CCPs are often prevented from keeping cash on commercial bank accounts, preferring to repo it out. Not surprisingly, care must be taken with the counterparties at which the CCP invests, attention paid to term length, and bonds offered in return, with concentration of counterparties and bonds.

5.4.9 Client Asset Protection

If a member defaults, a CCP protects its direct members and their clients from the effect. But the defaulting member's clients lose their positions and collateral. A well-known example is from Lehman Brothers in 2008 and MF Global in 2011.[18] This is true for both principal and agent models.

To resolve and broaden the counterparty credit risk management through to the second layer, CCPs have adopted various models for *client asset protection* or *segregation* as required by regulation (Exhibit 5.13).

This could reduce the profitability of client clearing. That's because of the loss of substantial revenue from collecting "gross" from clients and paying "net" towards the CCP (re-hypothecation) as well as the added legal and operational costs.

Final Word: Legally, a CCP is in some sense a super-senior counterparty in terms of the relevant insolvency regime. Upon a default, they are able to utilize their pledges, the CCP structure setting aside individual and mutual capital to reflect exposure from trading. In a default, it has transparent processes to mitigate and allocate losses. This has enormous macro-prudential implications, and is the primary reason for the mandated clearing promoted by regulators.

---

[17] Title transfer denotes the legal transfer of full ownership. In this case, the CCP becomes the owner of the cash.

[18] See, for example, https://www.newyorkfed.org/medialibrary/media/research/epr/2014/1412flem.pdf for the Lehman Brothers case, and https://www.chicagofed.org/~/media/publications/economic-perspectives/2015/4q2015-part1-ruffini-pdf.pdf for MF Global.

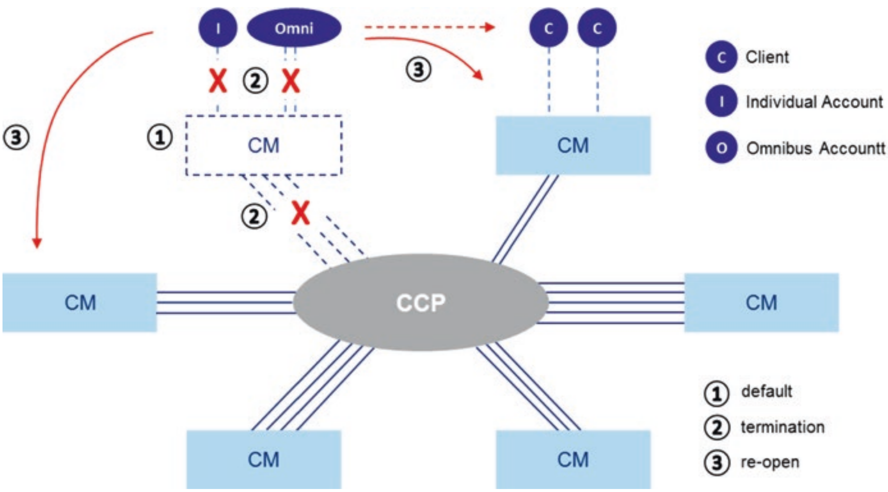


Exhibit 5.13 Segregated accounts and porting in default

5.5 CCP Structure and CCP Market Structure

The CCP is clearly valuable for the functioning of a market. The trade's counterparty is replaced for all members by the CCP, thus changing the credit risk profile of the trade. Centralized processing and value-added services of the CCP simplify and optimize clearing. New features, such as post-trade anonymity which can be beneficial to avoid procyclical cessation to trading can be incorporated only with a CCP. However, the central clearing also has some indirect effects, often relevant to the resilience of the market in crises and risk management. These features have led to the expansion of central clearing in the wake of financial crisis starting in 2007. The ripple effects of the crisis on CCPs are notable:

1. *Independent risk computations*. With CCP's incentives aligned to protecting the market in the case of default, it charges members conservative figures for the risk posed by their open trades.
2. *A neutral but strict mark-to-market process losses directly*. The end-of-day prices that CCPs employ for this process are the same across the market. All participants outstanding trades are evaluated transparently and consistently, an approach that ensures that losses do not accumulate between the trading parties.
3. *All trades cleared by the CCP benefit from the extremely high credit risk of the CCP*. The CCP frees up its members to devote their attention to the best price for the trade, and not the possible counterparty credit risk factors. That's because the credit quality of its members is replaced with the credit quality of the CCP, backed by its waterfall and default management process.

CCPs gain market share from more efficiency. This in turn has led to a limited number of CCPs for those asset classes in which the CCP's risk mitigation and multilateral netting benefits are superior.

5.5.1 Governance

The changes to the governance of CCPs have been substantial. During the course of the last two decades, most CCPs have been or were in the process of demutualization. Since the financial crisis, the focus on CCPs has had an impact on their governance, rooted in the desire for all stakeholders to have a more vested interest. The primary example is the risk committees of CCPs. These committers advise the CCP on risk matters such as initial margin levels for new products, waterfall structures, and liquidity management. EMIR, for instance, requires that CCPs have members and indirect participants on their risk committee.

5.5.2 Transparency

A key point about CCPs: they are an ex ante agreed-on mechanism to operate a rebalancing and from this a designed loss allocation if necessary along prearranged lines. Transparency is of paramount importance as this creates conditions for strong ex ante risk management standards from the participants. A CCP does not have own positions for which the disclosure might lead to gaming, but must be careful so as not to reveal client trading information to other players.

5.5.3 Accounting Practice of CCPs

Many of the preeminent CCPs have demutualized in the last quarter-century. As regular businesses, many of them public companies or subsidiaries thereof, they publish reports for their investors with market development data, such as traded and cleared volumes of contracts, nominal or notional as the case may be per asset class. The data also include risk mitigation quantifiers, namely, percentage of the above that was netted down, total resources in the CCP's waterfall, and CCP's capital.

Most CCPs apply either (US) GAAP or IFRS in their reporting, typically separating the business elements of the CCP—direct clearing fees and interest earned on collateral services, for example—from its role as trading counterparty to its members, collecting margin and clearing fund contributions and handling open positions.

A CCP's major assets and liabilities will always arise from the trades it processes and guarantees between its participants. These, however, are always exactly counterbalanced. The CCP simply stands between the actual beneficiaries to the trade until settlement is concluded. Additionally, the CCP's waterfall consists of substantial amounts of both cash and noncash assets on behalf of members. Cash collateral is virtually always held as title transfer and noncash collateral—most commonly high-quality bonds—as pledges. The CCP becomes unbalanced in the event of a member default. In this case, the CCP also has the right to use the collateral of the defaulter as well as potentially non-defaulters' assets for temporary liquidity purposes.

5.5.4 Business Model: Costs

There are three types of CCPs: member utility, government run, and private. The member-owned[19] utility model has decreased in popularity. The challenge of a publicly run CCP is that the counterparty credit risk of the market participants is effectively underwritten by the public. Private CCPs, on the other hand, must provide greater skin in the game to inspire confidence from their members.

5.6 Outlook

5.6.1 The CCP Model Under Review

Since the crisis, there is a keen interest in the contribution and resilience of CCPs. A global regulatory review showed that CCPs performed very well in the limited number of defaults. However, regulation in the area of CCP risk management had been relatively light and heterogeneous. Despite their excellent performance, CCP stakeholders are keen to ensure that all CCPs are prudently managed to the highest risk management standards. This is especially relevant for the "new" markets where CCPs are expanding, an expansion driven both by the markets themselves and by regulatory instigation.

An example of the former is exchange-traded derivatives; the latter occurs either directly through capital incentives (Basel) or by fiat (clearing obligation). Indeed, all major jurisdictions have introduced legislation and enacted laws on the specifications of a CCP. The regulatory backdrop features a "qualified CCP" (Basel, CPSS/IOSCO), outlining the basic functions of a CCP, as well as capital treatment from the member's perspective. An important development is the CCP capital and risk management structure. Previously, most CCPs were a "good till the last drop" CCP, in which losses exceeding the prefunded waterfall were called from members. That was often in rounds of size equivalent to the initial contribu-

[19] Many CCPs were, and some still are, jointly owned by their members. Each member often holds shares in the CCP based on their volume at the CCP.

tion. Members unwilling or able to provide further contributions retained the right to exist the CCP service. In recent years, however, members have pushed for "limited liability" towards CCPs, so that they are only exposed to a fixed, finite number of replenishments. The effect is that CCPs may run out of funds in a truly dramatic crisis. In response, recovery and resolution plans are drafted for CCPs, outlining that CCPs will have to ensure that losses can be contained and absorbed without recourse to public funds.

It has been stated that CCPs are "super-systemic." They form, to a large degree, a system itself, any disruptions at the CCP directly affecting all of its members. Some have claimed the title of "systemic risk managers" for CCPs, seen as a mechanism for a market to address any disruptions. The truth probably is in between because CCPs basically have a stable business model, based on a low-risk appetite with much of the actual collateral acquired from the defaulter pays/mutualization construct. On the other hand, stringent operational risk criteria must be followed so that CCPs can serve their markets well.

All things considered, CCPs have weathered financial market shocks and member defaults as designed.

That said, in a handful of cases, their risk management has been inadequate, or else the market conditions overwhelmed their safety nets. In addition to the near misses and liquidity issues in Chicago in 1987, the primary cases of CCP disruptions occurred at the Caisse de Liquidation (Paris, 1974), The Kuala Lumpur Commodities Clearing House (1983), and the Hong Kong Futures Exchange (1987). In all cases, a combination of unchecked risk buildup at certain members, followed by a rapid drop in prices, put the CCP under strain.

The Black Monday-related market drop in Hong Kong has been thoroughly documented as part of a governance and technical review of the CCP following a joint government, member, and shareholder emergency loan to the CCP. In essence, this was the first recovery plan operated for a modern CCP. The lessons in terms of charging appropriate concentration limits and guarantee fund sizing still have impact today. The case for implementing the highest risk management standards possible for CCPs globally—for instance, with EMIR (European Market Infrastructure Regulation) in Europe, or Dodd-Frank in the USA—comes from carefully analyzing these incidents.

One challenge in the future is that CCPs will tend to have very similar risk frameworks and models. This means that any problem with the common elements poses a risk to all CCPs, since homogenous risk management is exposed to the same problems.

Henceforth, it is unclear what the best role model is for a CCP. It is, indeed, an open question.

At present, the preeminent CCPs combine aspects of a private company that could suffer non-default losses, a mechanism for the market to resolve credit events and their spillover, and a liquidity provision hub. CCP competition and choice for market participants so far are drivers of innovation and for strengthening operating standards.

The clearing obligation raises another question, driven by considering DMP alternatives for large directional positions from buy-side entities such as pension funds, insurance companies, or general corporates. For these, the degree of mutualization, and even the assumed 100 % probability of default by CCPs across all participants, might be revised. This point is more poignant for state-backed entities or lower leverage financial actors that are signing up to CCP services in repo and swap markets.

5.7 Final Word

These are exciting times for CCPs: they are in the spotlight like never before, adapting to new regulations and clearing mandates. Their systemic contribution to the financial system, previously obscure or relegated to a limited number of experts, receives greater attention. In the future, the risk mitigation tools of CCPs in the next systemic disruptions will be critical—and every effort should be made so that CCPs are not "fighting the last crisis."

Acknowledgments The author would like to acknowledge the valuable contribution of Mr. Teo Floor and thank him for his continued support.

Further Reading

Eurex Clearing / Deutsche Börse Group (2014): How central counterparties strengthen the safety and integrity of financial markets. Frankfurt / Main: Deutsche Börse AG.

Robert Schwartz, Reto Francioni. Equity Markets in Action, 2004

Peter Norman. The Risk Controllers: Central Counterparty Clearing in Globalised Financial Markets. 2011

Paul Tucker: Clearing houses as system risk managers: http://www.bis.org/review/r110608g.pdf

Jon Gregory. Central Counterparties: Mandatory Central Clearing and Initial Margin Requirements for OTC Derivatives. 2014

CPSS IOSCO Principles for Financial Market Infrastructure: http://www.bis.org/cpmi/info_pfmi.htm

European Market Infrastructure Regulation: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32012R0648

The Dodd-Frank Wall Street Reform and Consumer Protection Act: http://www.cftc.gov/idc/groups/public/@swaps/documents/file/hr4173_enrolledbill.pdf

Darrel Duffie. Selected papers: http://www.darrellduffie.com/

Chapter 6
Securities Services: Settlement, Custody and Financing

Marc Robert-Nicoud

6.1 Introduction

Securities services, an essential and constant part of the capital market industry, have historically escaped the limelight. Securities service infrastructure has shown its resilience—and sheer importance—in the face of market upheaval during the financial crisis of 2008. And yet, in many important ways, securities services have recently taken center stage in the regulatory environment. Why this sudden and acute regulatory interest today?

Willie Sutton, the famous American bank robber of the 1930s, when asked why he robbed *only* banks, reportedly said, "That's where the money is!" The securities service infrastructure, in a similar way, is ultimately where the financial assets "are," and where these assets change hands. It must be safe, and it must be efficient.

The regulatory response to the last crisis has seen significant investments channeled into technologies and processes that are designed to reduce risks. These developments can also be leveraged to reduce banks' operating costs. Market harmonization initiatives like TARGET2-Securities (T2S)[1] and regulatory requirements add unprecedented momentum to the evolution of securities services.

This chapter starts with a brief overview of the fundamentals of securities services. We then turn to a review of the types of risks faced by market users, and to the unique role of securities service providers in mitigating these risks. A review of securities services follows, starting with the issuance of new securities, then settlement services with a focus on delivery-versus-payment models, and how settlement finality is achieved. Settlement naturally leads to securities financing, which is a driver of efficiencies and risk reduction. Securities under custody, whether used as

[1] See Sect. 6.5.4 for further details.

M. Robert-Nicoud (✉)
Clearstream Holding AG, 61 Mergenthalerallee, Eschborn 65760, Germany
e-mail: Marc.robert-nicoud@clearstream.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_6

collateral, or as the object of a transaction, require servicing in relation to income payments and corporate events. An exploration of these custody services completes the review.

6.2 Basics

6.2.1 The Nature of the Business

Securities services are a natural consequence of any trade execution. This is the case regardless of whether the parties trade on an exchange, or over the counter (OTC); whether domestic or international securities; or whether the trade is part of the primary (new issuance) market or the secondary market. The most common securities are equities (shares) and fixed-income securities (e.g., bonds). Equities entail ownership in a company, and they document the rights associated with ownership. Fixed-income securities certify the right to obtain interest and redemption at maturity. Both bonds and equity trades require comparable securities services. The majority of equity trades take place on exchange; fixed-income instruments primarily trade OTC. Figure 6.1 shows the evolution of securities services from the traditional, such as settlement and safekeeping, to complementary value-added in markets around the world.

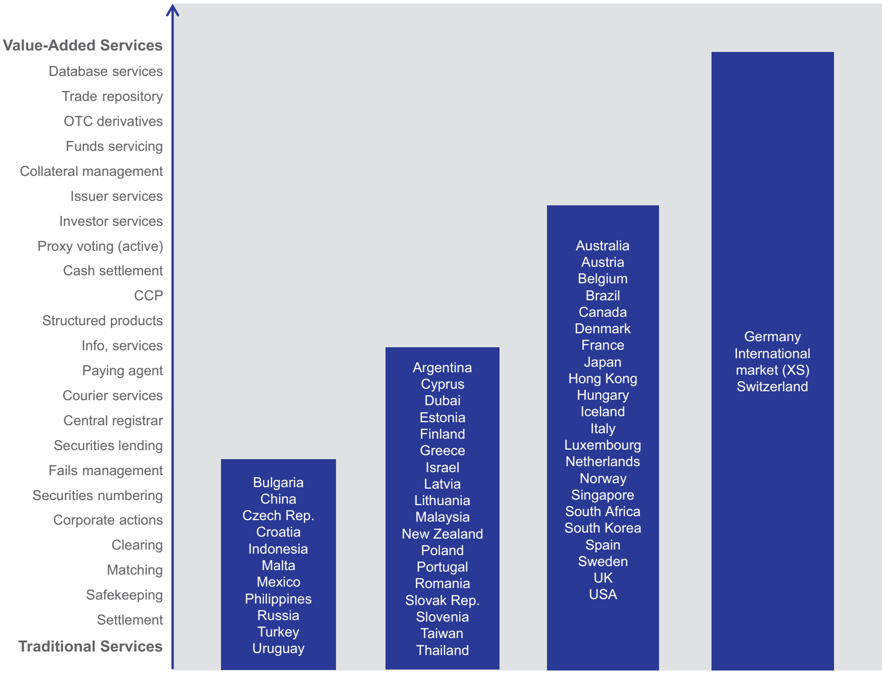


Fig. 6.1 Development of value-added securities services across markets (Thomas Murray Data Services: CMI in Focus - Value added services by CSDs, 2014. Available at: ds.thomasmurray.com/opinion/cmi-focus-value-added-services-csds)

For orientation, it is useful to connect some of these services together at the outset. Settlement services are required to execute the transfer of ownership and cash after a securities trade. Settlement can have several purposes: mainly a sale or purchase, but also the transfer of securities and cash in the context of primary issuance, collateral management, and securities lending. At any point during the life of a security, it will have to be safely kept (physically or electronically). The corporate events relating to the security will call for corporate action processing, also referred to as asset servicing. Safekeeping and corporate action processing together can be referred to as custody services. A number of ancillary banking services exist in conjunction with traditional securities services, specifically credit and cash management services.

At the heart of this chapter is the concept of a "security." The concept can have different definitions in different jurisdictions. This chapter adopts a functional approach that side-steps this mostly legal debate. The definition adopted by UNIDROIT (an intergovernmental organization based in Rome that focuses on the harmonization of private law between states), in its convention on substantive rules for intermediated securities, allows us to focus on this essential: "any shares, bonds or other financial instruments or financial assets (other than cash) which are capable of being credited to a securities account and of being acquired and disposed of in accordance with the provisions of this Convention."[2] In Sect. 6.4, the most common types of securities are described as part of the securities issuance process.

In summary, securities services exist to support securities markets across a wide range of areas, from the issuance of securities to their redemption or divesture. That latter area includes processing the rights and obligations that result from owning and transferring securities. Put simply, securities services are diverse, with some being at the core of the financial infrastructure.

Figure 6.2 shows where securities service providers are embedded in the financial market value chain. As elaborated on later in this chapter, a well-functioning securities market relies on securities services and on the accompanying providers to ensure issuer (sell-side) and investor (buy-side) protection. Securities service providers contribute to creating the best conditions for the supply and demand of the securities markets. Not surprisingly, well-functioning securities services are essential to stable, fair, and performing financial markets, given the importance of the securities market to the financial industry.

6.2.2 Market-Structure Development

Financial infrastructure is, in essence, local in nature. But in terms of reach, it is as global and far flung as the markets it serves. The most relevant securities markets are in the USA and Europe, and then there is also the international (Eurobond) market. Europe is the most sophisticated, and the most active cross-border market. Figures 6.3 and 6.4 show the level of internationalization in European and US debt holding structure.

---

[2] UNIDROIT, 2009. "Unidroit convention on substantive rules for intermediated securities." Available at http://www.unidroit.org/english/conventions/2009intermediatedsecurities/convention.pdf.

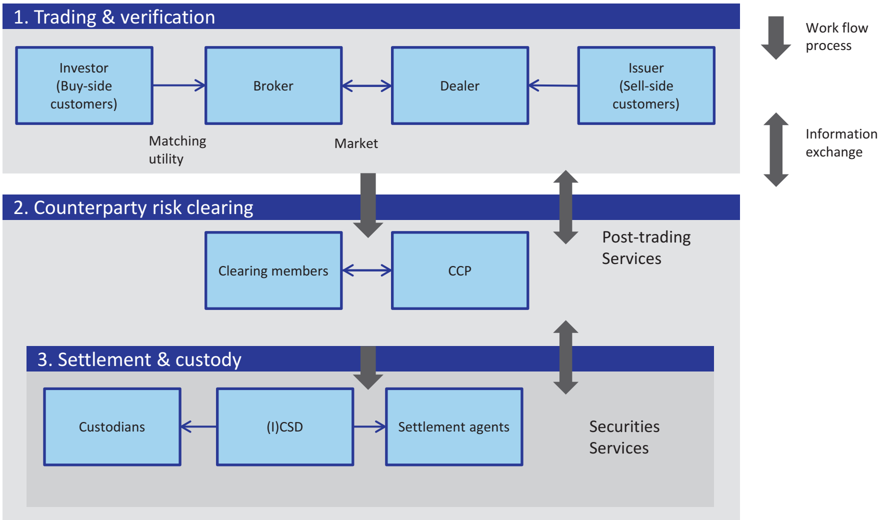


Fig. 6.2 Stylized illustration of the value chain and securities services

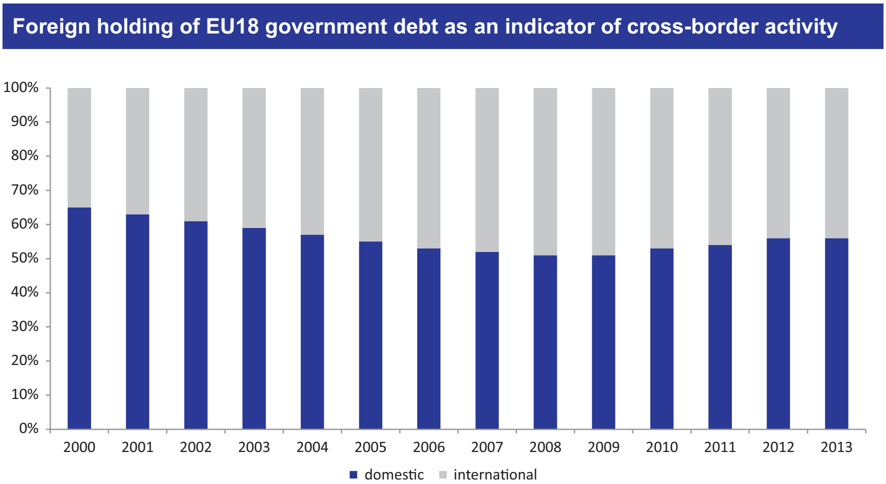


Fig. 6.3 Foreign holding of government debt as an indicator of cross-border activity (Accenture Research, 2014)

The European securities market is also among the most dynamic, undergoing a complete transformation of its settlement infrastructure with T2S. This chapter consequently pays greater attention to the European landscape while, at the same time, making frequent reference to the USA and the international market.

Securities services are tailored to the applicable legal environment, market practices, regulation, and taxes. Many of these factors weigh in favor of local specificities. Significantly among them is investors' preference for local investment, a preference sometimes referred to as a "home-bias."

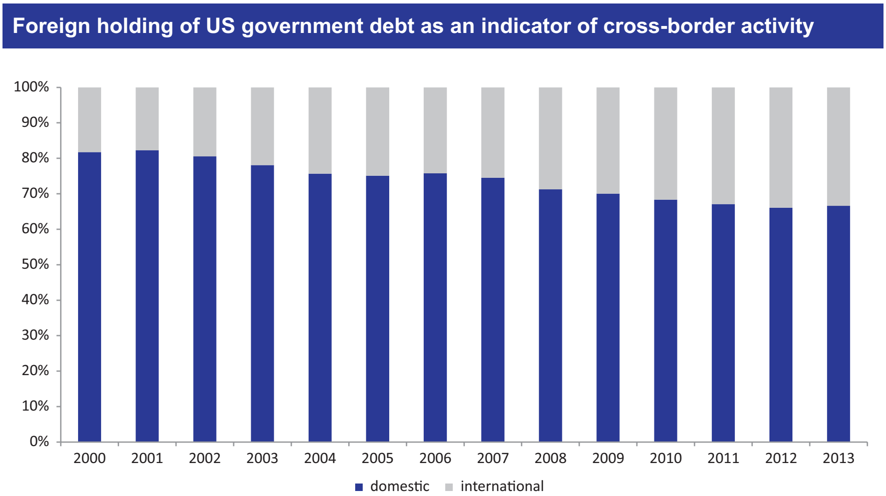


Fig. 6.4 Foreign holding of government debt as an indicator of cross-border activity (Bureau of the Fiscal Service, Department of the Treasury, 2014. "Treasury Bulletin." Available at http://www.fiscal.treasury.gov/fsreports/rpt/treasBulletin/b2014_4.pdf)

The European Union (EU), and especially the Eurozone, has pushed for a harmonization of trading, clearing, and securities services aiming to leverage a single market for securities. The UN, similarly, has sponsored global harmonization initiatives such as UNIDROIT's Convention on substantive rules for intermediated securities. A progressive harmonization of laws and regulation will certainly support the creation of regional, or even global, securities markets.

Harmonization also has a cost dimension. Securities services are generally delivered according to a scalable business model, and large investments in infrastructure are needed for securities services that satisfy market and regulatory standards. Average costs tend to be inversely related to volume: the larger the volumes, the lower the costs. For instance, large markets tend to have a lower cost of settlement than small ones, see Fig. 6.5. Harmonization between markets is a necessary first step to volume consolidation and the prospect for cost reduction that follows.

6.2.3 Securities Service Providers

The securities service industry is supported by a sophisticated intermediation structure to bring investors, issuers, and, more generally, trading parties together in the performance of specialized and complex processes. Numerous providers offer securities services—from central securities depositories (CSDs) and international central securities depositories (ICSDs), to custodians and other banks. These institutions compete for securities services although they do not necessarily have the same service scope, as shown in Fig. 6.6.

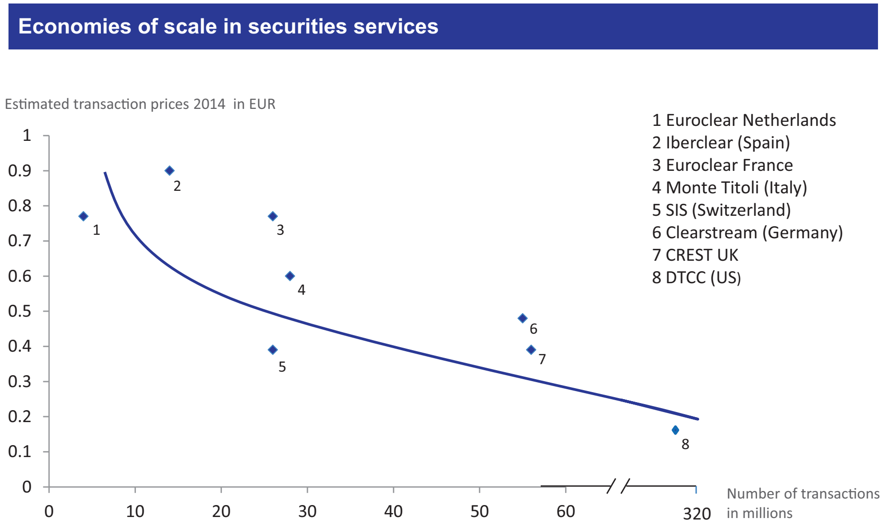


Fig. 6.5 Economies of scale in securities services

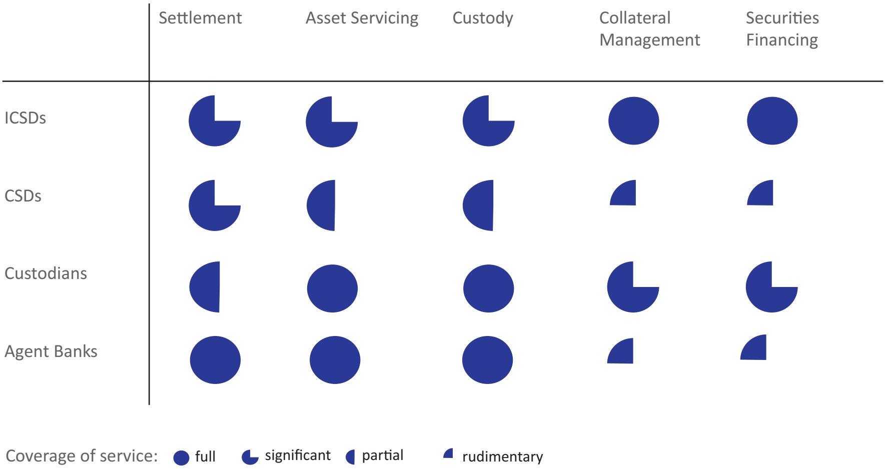


Fig. 6.6 Varying service scopes of securities service providers (Clearstream Research and Data, 2014)

The services are described later in this chapter. Notably, most of the institutions mentioned in this chapter are directly or indirectly connected to each other by the services they offer or receive. The very competitive nature of the securities service space does not diminish the partially complementary nature of the various offerings, as well as from their importance for the reliability of financial markets.

A short, simplified overview of some typical interaction between these securities service providers is useful to put in context the information later in this chapter. The interaction between these entities is best understood by following the settlement and custody process chains:

- Custody/security holding chain: An investor holds a position in a security. The investor may have physical possession of the security, or may have the position in deposit with a bank. The bank, in turn, can either have the securities in its own vaults, with a custodian, or with a central securities depository. The nature of the rights in a custody chain is a function of the legal jurisdiction of the holding, a matter that can sometimes lead to complexity in the case of cross-border holdings. Take the example of a French investor holding a US security in a French bank account. The bank may hold the position in the US security with a global custodian and its US sub-custodian, with an ICSD and its sub-custodian, with a local US custodian, or directly with the relevant US CSD. (The Federal Reserve Banks act as CSD for all marketable US Treasury securities, and the Depository Trust and Clearing Corporation or DTCC for other securities.) The choice is generally not the investor's, but lies with the investor's bank. The decision will be a function of cost, quality of access (e.g., market deadlines), relationship, counterparty location, and securities financing services.
- Settlement/security transfer chain: An investor seeking to buy a position in a security will likely leverage his or her retail bank to execute and settle a trade. The bank will either settle internally (if the counterparty uses the same bank) or will use a settlement agent or a custodian (if the counterparty uses a different bank). In the latter case, the two custodian banks will transfer the securities and the cash in a central securities depository. For the purpose of this settlement, the CSD also acts as a securities settlement system (SSS), and it will utilize a payment system. Depending on the security and the market, the trade may also be reported in a trade repository.

The term "custodian" typically refers to entities with different business models and service scopes. The services they offer are, however, comparable. In fact, they can also be compared to the services of CSDs and ICSDs (Fig. 6.7).

6.3 Review of Risks mitigated by Securities Service Providers

Securities services are generally reliable and safe. In fact, these services help market users to reduce and mitigate their risks. Safe and efficient securities service providers actually help reduce systemic risk. But, of course, providers are not entirely immune from risk. However, most service providers in this field are strictly regulated, and they adopt very low risk profiles. This is justified because their inability to perform could have significant and adverse effects on the markets they serve, and on the

Term	Description
Custodian	An organisation that holds securities and (usually) cash on its clients' behalf; and may effect settlement of trades on its clients' behalf. Custodians can also be split into varying degrees of scale: - Global Custodians: custodians that safekeep assets for their clients in multiple jurisdictions around the world - Sub-Custodian: a custodian within a global custodian's network of custodians
Central Securities Depository (CSD)	An organisation that holds securities, normally in *book-entry* form; usually the ultimate place of settlement, effected through book-entry transfer
International Central Securities Depository (ICSD)	A CSD that handles domestic and international securities. Only two organisation are recognised as ICSDs, namely Clearstream Banking and Euroclear Bank.
Settlement agent	An organisation that effects the exchange of securities and cash on behalf of ist clients; resultant securities and cash balances may or may not be held

Fig. 6.7 Definitions of certain securities service providers (Simmons, M.: *Securities Operations: A Guide to Trade and Position Management*, Wiley, 2002, p. 227)

broader economy in the event of an upheaval. Securities service providers that play an infrastructure role like CSDs and ICSDs typically do not assume principal risk in the execution of their customers' instructions (Fig. 6.8).

The importance of risk management in current financial policy initiatives calls for the topic to be discussed early in this chapter. Risks that are of relevance to securities services can be categorized as credit risk, liquidity risk, operational risk, and legal risk. Figure 6.9 outlines these risk categories, followed by a brief examination of how each impacts and is mitigated by market infrastructure such as CSDs and ICSDs. It is worth remembering these risks throughout the securities service descriptions that follow this section.

6.3.1 Credit Risk

Credit risk is the risk that a counterparty—that is, a participant or other entity—will be unable to fully satisfy its outstanding financial obligations at any time, or will otherwise fail to honor the terms and conditions of an agreement. Credit risk is inherent in all activities that depend on the counterparty's, issuer's, or borrower's performance. This risk typically arises each time funds are extended,

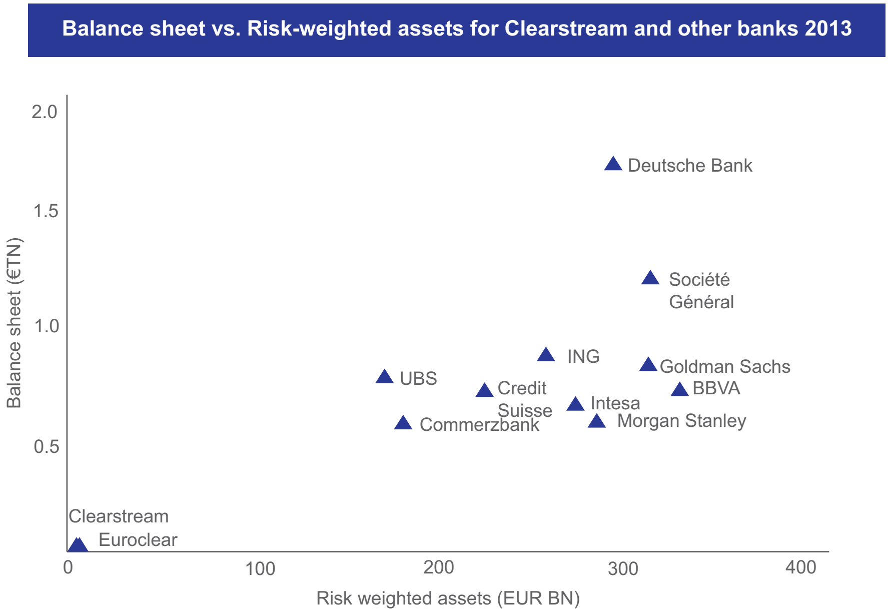


Fig. 6.8 Balance sheet vs. RWAs for ICSDs and various banks, 2013 (Clearstream Research and Data, 2014)

committed, invested, or otherwise exposed through actual or implied contractual agreements. Indeed, any trading party is subject to credit risk because the counterparty may default before delivering the cash or securities. If there is a failure to deliver cash or securities, additional costs may be incurred by the parties of the transactions.

In a number of ways, the services of securities providers help to reduce credit risk considerably. For instance, there is a risk of delivering securities without receiving payment, or paying without receiving securities in any trade. Many (I)CSDs in their capacity as SSS perform settlement in a delivery-versus-payment (DVP) arrangement. This means that the exchange of cash and securities can occur practically simultaneously in the books of the SSS.

In addition, many SSSs in Europe offer settlement in central bank money, the cash leg taking place in the books of a central bank. The risk of cash not being available is significantly reduced in this way. SSSs also typically apply a settlement finality rule, which mitigates the risk. Once final, a settlement cannot be reversed for the benefit of the creditors of a counterparty. In the collateral management area, some (I)CSDs offer tri-party repo services, as part of which the exchange of securities and cash also take place in DVP mode in the safe environment of a CSD. In most jurisdictions, CSDs do not own the securities of their customers, not even on their behalf. Consequently, if they default, securities cannot be used to satisfy any debt obligations of the CSDs.

	Source	Example causes	Emerging complexity drivers
Credit risk	Default by borrower, leading to loss of principal and/or interest	• Credit default risk • Country risk	• Agency models • Bilateral margining • Shift to central clearing
Liquidity risk	Lack of funds to support operations or inability to exit market position	• Funding liquidity risk • Market liquidity risk	• Withdrawal of repo capacity • Execution venues for new asset classes
Operational risk	Failure of internal processes, people and systems	• Process failures • Systems failures/IT risk • Damage to physical assets • Compliance/legal risk	• System outages • Interconnectedness • New technologies
Legal risk	Unexpected application of law or regulation	• Uncertainty of laws • Conflicting bodies of law	• Cross-border transactions

Fig. 6.9 Outline of four categories of risk (Oliver Wyman, 2014. "The Capital Markets Industry: The times they are a-changin'." Available at http://www.oliverwyman.de/media/The_Capital_Markets_Industry_-_The_Times_They_are_A-Changin.pdf)

6.3.2 Liquidity Risk

Liquidity risk is the risk that a counterparty will have insufficient funds to constantly satisfy its financial obligations anytime expected, notwithstanding the counterparty's ability to potentially do so in the future. Liquidity risk is the risk that the seller of an asset will not receive payment when due, and the seller may have to borrow or liquidate assets to execute other payments. It also includes the risk that the buyer of an asset will not receive delivery when due, the buyer having to potentially borrow the asset in order to execute onward delivery obligation. Thus, both parties to a financial transaction are exposed to potential liquidity risk on settlement date.

(I)CSDs and SSSs contribute to the mitigation of this risk by offering emergency borrowing services to lend funds or securities, for short periods, in anticipation of receiving delayed cash or securities. Generally, securities service providers have a very positive impact on market liquidity. That is because the efficiencies they create enable assets to move as part of faster and yet safer processes.

6.3.3 Operational Risk

Operational risk is the risk that deficiencies in information systems or internal processes, human errors, management failures, or disruptions from external events will result in the reduction, deterioration, or breakdown of services. These operational failures may lead to delays, losses, and liquidity problems. This risk is inherently present in all human activities. Both users and providers of securities services must continuously manage operational risk. There are some additional risk factors: different rules and conventions by markets; different types of securities, capital, or currency restrictions; availability and communication of timely and accurate information; and degree of automation in different markets. The automation of securities services in this field helps to mitigate the risk. But experience shows that errors in corporate action and settlement processing are common causes of losses.

These errors can be greatly mitigated by effective risk identification and control. Moreover, effective policies and procedures, a strong control environment, and efficient use of technology are essential risk management tools. Automation and use of straight-through processing (STP) significantly reduce operational risk.

6.3.4 Legal Risk

Legal risk is the risk of the unexpected application of a law or regulation, usually resulting in a financial loss. Legal risk can also arise if the application of relevant laws and regulations is uncertain. For example, legal risk encompasses the risk a counterparty faces from an unexpected application of a law that renders a contract illegal or unenforceable. Legal risk also includes the risk of loss resulting from a delay in the recovery of financial assets, or a freezing of positions resulting from a legal procedure.

In cross-border as well as some national contexts, different bodies of law can apply to a single transaction, activity, or participant. In such instances, the infrastructure and expertise provided by securities service providers substantially reduce the risk on market users. In addition, the depth of experience and expertise of securities service providers allow them to conduct detailed research into the impacts of regulatory initiatives, a matter that is useful for informing the legislative process.[3]

6.3.5 Conclusion of Risk Review

The most likely risk to materialize in securities services is operational risk. Still, the most prominent risks mentioned in industry circles are the credit and liquidity risks—probably because they are more closely associated with systemic risk. Put

[3] For example the standards developed by the International Securities Market Advisory Group, see the website of the International Capital Market Association for more. www.icmagroup.org.

differently, as the Committee on Payments and Market Infrastructures (CPMI) describes in its Principles for Financial Market Infrastructure, CSDs and other market infrastructure may face systemic risk caused by the inability of one or more participants in a market to perform as expected. Consequently, "knock-on" effects are possible.[4] Moreover, the complex links between market infrastructure can also cause disruptions passed from one market to another. Clearly, an inability of market infrastructure to complete settlement could have significant adverse effects on the markets it serves.

On the other hand, market infrastructure such as CSDs, in their function as SSS entities, contribute significantly to the mitigation of credit and liquidity risk. They do so by offering DVP settlement and by applying settlement finality rules, such as the Settlement Finality Directive in Europe.

As agents, securities service providers are not directly exposed to credit and liquidity risks. (I)CSDs providing ancillary banking services assume marginal principal risk, and are tightly regulated by prudential supervision. They are not party to the transactions they settle for their customers. The securities held by the (I)CSDs on behalf of their customers do not appear on their balance sheet since they remain legally owned by their customers.

There is no evidence, nor any suggestion, that CSDs contributed to the financial crisis of 2008. And many analysts would argue that theirs was a positive role: providing liquidity and safety when the markets were lacking both. Regardless, the critical part that CSDs play as securities settlement systems and as custodians has made a higher degree of scrutiny a legitimate issue.

In the aftermath of the financial crisis, regulators have put pressure on banks to ensure that they have credible recovery plans that do not require extraordinary governmental support. And they have put pressure on banks to also have "living wills"—plans for the orderly wind-down of their services. Market infrastructure is not immune to this requirement that the collapse of a single institution does not endanger the entire market. That is clear from reports such as the CPMI-IOSCO study of October 2014 on the Recovery of Financial Market Infrastructures.[5]

Pursuant to the report, the purpose of a recovery plan is the identification of the FMI's critical services, stress scenarios, and recovery triggers, as well as a substantive description of its recovery tools and tools to address structural weaknesses.[6] In essence, the report argues that by maintaining structured and realistic plans for covering these topics, market infrastructure are not just introducing steps to maintain their own resilience under extreme circumstances. More generally, under extreme circumstances, the steps also benefit markets that are reliant on these infrastructure services.

---

[4] Committee on Payments and Market Infrastructures, 2012. "Principles for Financial Market Infrastructures." Available at http://www.bis.org/cpmi/publ/d101.htm. N.b.: the Committee on Payment and Settlement Systems (CPSS) changed its name to the Committee on Payments and Market Infrastructures (CPMI), on 1 September 2014.

[5] Committee on Payments and Market Infrastructures, 2014. "Recovery of financial market infrastructures." Available at http://www.bis.org/cpmi/publ/d121.pdf.

[6] *ibid*, pp. 8–11.

6.4 Issuing Securities

Companies and governmental agencies seeking sources of financing can either ask banks for a loan or turn to the financial markets directly to raise capital. Those who raise capital through the creation and distribution of securities to investors are known as issuers. They can either sell part ownership in the company represented by shares or borrow cash from investors through bonds. Government entities, whether local or supranational organizations, can only issue bonds. In contrast, companies or corporations also issue shares (equity).

The majority of trades executed in regulated securities marketplaces around the globe on a daily basis are in outstanding securities (i.e., securities that have already been issued). Securities in this category are traded in the secondary market. In contrast, newly issued securities are traded and settled in the primary market (Fig. 6.10).

6.4.1 Methods of Issuing Securities

Bringing securities to the marketplace requires special skills to manage the process efficiently, an expertise an issuer of the securities (either a company or a government) does not generally possess. It is therefore common practice for issuers to appoint specialist agents who typically belong to the corporate banking and investment

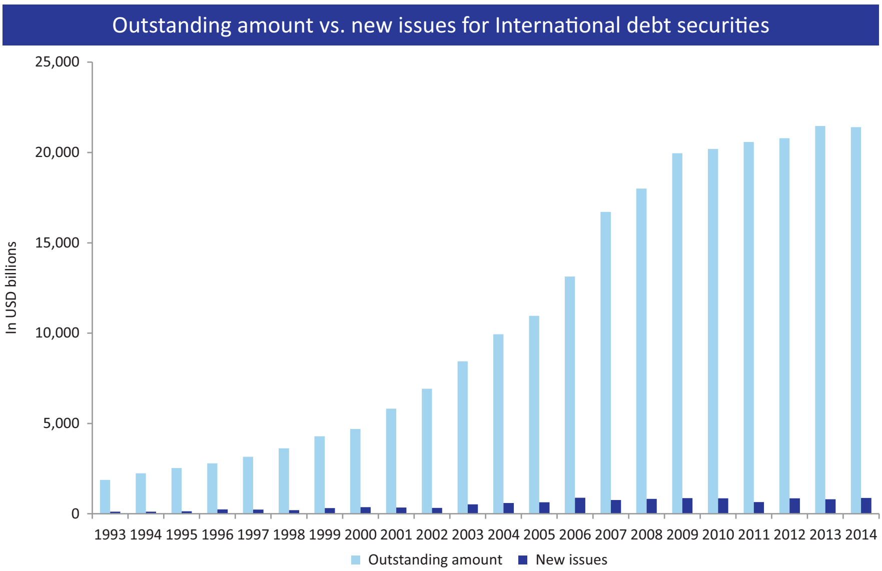


Fig. 6.10 Outstanding amount vs. new issues for international debt securities (bonds and notes) (Bank for International Settlement, Debt Securities Statistics. Available at https://www.bis.org/statistics/secstats.htm)

banking areas. These specialists coordinate the activities of numerous agents, from the calculation agent to the paying agent and the legal advisors to others. They will advise issuers, distribute the securities to investors, and ensure that the issuer receives the cash at the agreed time.

Securities can be issued by various methods for equity or debt, based on the market conventions where the issue is launched. In most cases, the issuer will publish a paper, sometimes called a prospectus or term sheet indenture, detailing the quantity of the issue and its price, as well as a description of the issuing company.

Bonds can be issued via either auction or syndication. In an auction, securities are sold to the highest bidders by, for example, central banks or funding agencies on behalf of their governments. In a syndication, the lead manager creates a consortium of banks and other market participants who work together to promote the distribution of the bond to investors.

Shares are usually issued via an initial public offering (IPO), advertised in the media and by a prospectus. A syndicate of market participants receives share allotments from the lead manager (on behalf of the issuer), and then allots some or all of these shares to their clients. Shares can also be issued in a restricted manner in a private placement by only offering the shares to a select group of institutional investors, among them pension funds and insurance companies.

Securities can be issued in registered or bearer form. The issuer or its agent keeps records of the holders in registered securities. Securities issues in bearer form are, theoretically, traded without any record of ownership, so physical possession of the security is the sole evidence of ownership. In practice, the majority of securities issued to the public, bearer or registered, are now immobilized and represented by electronic book entries in which the transfer of securities ownership is realized by crediting or debiting the seller's or purchaser's account without transferring physical certificates between counterparties.

6.4.2 Issuing via (International) Central Securities Depositories

Lead managers often turn to (I)CSDs to bring a security issue to the market because they possess the critical infrastructure for distribution, settlement, and safekeeping. Some (I)CSDs also provide value-added services such as the custody services described later in this chapter. That means that they can service a security throughout its entire lifecycle.

Clearstream Banking S.A., for example, offers the issuing community a wide range of products and services for debt, equities, investment funds, warrants, and structured products. Moreover, it provides the infrastructure for issuers to reach investors anywhere in the world.

Services offered by (I)CSDs include eligibility assessments, issuance, and distribution of domestic, foreign, and international new issues of global and domestic instruments: certificates of deposit, depository receipts, treasury bills, commercial papers, short-term and medium-term notes, bonds, equities, warrants, equity-linked notes, and investment fund shares.

(I)CSDs can also assist lead managers, lawyers, and issuing agents by reviewing issue structures. They can also provide additional information on operational procedures and the applicable documentation in the offering memorandum, prospectus, and agency agreement.

In the case of bond issuance, we can distinguish between securities issued to a domestic market, securities issued to a foreign market, and so-called international security types such as "Eurobonds."

6.4.2.1 Domestic Bonds

Domestic bonds are issued by resident borrowers in their own local market and currency, and held by their local CSD. An example is a US issuer issuing in USD via DTCC in the USA.

6.4.2.2 Foreign Bonds

Foreign bonds are securities issued in domestic markets by borrowers domiciled in foreign countries. They are normally denominated in the currency of their issuing market. These securities are held in the CSD of the domestic market where they are issued, and are often assigned colloquial names reflecting these (foreign) domestic markets (Fig. 6.11).

6.4.2.3 "Eurobonds"

CSDs handle domestic securities. ICSDs, on the other hand, are specialized in servicing foreign and international debt securities, known as Eurobonds. Peter Norman defines a Eurobond as "issued by a borrower outside its own country that may be denominated in a currency foreign to the borrower or to the purchaser, or both, and that is not subject to withholding tax or other legislation by the host country, in whose currency the bond is issued."[7] An example is a US issuer who might decide to issue bonds under UK law denominated in Euros to make its bonds more attractive to foreign investors.

---

[7] P. Norman, *Plumbers and Visionaries—Securities Settlement and Europe's Financial Market*, Wiley, 2007, p. 316.

Colloquial name	Currency	Issuer	Market
Yankees	USD	Non-US domiciled	USA
Bulldogs	GBP	Non-UK domiciled	UK
Matadors	EUR	Non-Spanish domiciled	Spain
Samurais	JPY	Non-Japanese domiciled	Japan
Pandas	CNY	Non-Chinese domiciled	China

Fig. 6.11 Common colloquial names for certain foreign bonds

The features of Eurobonds have resulted in ICSDs developing extensive distribution networks. As a result, this enables them to distribute new issues across multiple jurisdictions in the vast majority of currency denominations not subject to restrictions.

6.4.3 Identification of Securities

The International Securities Identification Number (ISIN) is the only internationally recognized standard for the unique identification of a security. The international reach is vast. ISIN is a key data field for cross-border trading that is used globally to identify securities, and it is recognized by many regulators as a mandatory data field for transaction reporting. When a CSD initially accepts a security, it is provided with an ISIN code and a Classification of Financial Instruments (CFI) code to permit its easy identification throughout its lifetime. The unique identification of a security is essential to the orderly trading and post-trading of any security. The rules governing the allocation of ISINs are set by the International Organization for Standardization (ISO) and the Association of National Numbering Agencies (ANNA).

An ISIN consists of three parts: a two-letter country code, a nine-character alphanumeric national security identifier, and a single check digit. For example, SIX Financial Information is the official securities numbering agency in Switzerland (prefix CH), Liechtenstein (prefix LI), and Belgium (prefix BE). To this end, it provides ISINs for equities, bonds, and other financial instruments issued in Switzerland. US ISINs are allocated by CUSIP Global Services (CGS), which is managed on behalf of the American Bankers Association by S&P Capital IQ. ISINs for Eurobonds consist of the nine-digit code common to both ICSDs for the issue, preceded by “XS” and followed by a numeric check digit. ANNA provides complete and updated information on ISINs and numbering agencies. The CFI code allows the grouping of financial instruments in a consistent way throughout the industry compatible with ISO rules.

European Pre-issuance Messaging System
Clearstream Banking, Euroclear Bank, and Depository Trust and Clearing Corporation (DTCC) have a joint initiative to increase the speed and efficiency of ISIN and common code allocation for selected money market instruments. Launched in 2002, the European Pre-issuance Messaging System (EPIM) platform is an automated, secure system that uses standard messaging formats and a standard messaging protocol to disseminate issuance information between the relevant primary market participants. For more information, please visit www.clearstream.com/epim.

6.4.4 Closing and Distribution

Once an issuance is closed, i.e. terms and allotment are agreed, payment is made to the issuer and the securities are allotted to investors. After a CSD or an ICSD (the typical place of a primary deposit) confirms receipt of the security from the issuer, they distribute their respective holding to the accounts of investors to whom the securities are to be allotted. Distribution to the issuer or his or her agent is then confirmed.

All subsequent trades of these securities will now take place in the secondary market, regardless of whether owners choose to sell the securities immediately, or to hold them. The secondary market for bonds is limited in that it comes to a close with the maturity of the bond. The secondary market for equities ceases only as a result of an event that causes the shares to no longer exist (i.e. a takeover, or a company bankruptcy).

6.5 Settling Transactions

6.5.1 Settlement and Finality

A number of securities service firms provide settlement services to their customers. At banks, this usually refers to contractual settlement (a contractual right against the bank which, in turn, will have to secure a true settlement), that is, unless the transaction is internalized by the bank. At this point, the primary market for the security no longer exists. This is because the securities are brought to the marketplace and the issuer has received cash in exchange for them (Fig. 6.12).

Ultimately only SSSs provide final settlement to effect the definitive transfer of property between buyers and sellers of intermediated securities. SSSs ensure and clearly define the moments of enforceability and irrevocability of transfer orders,

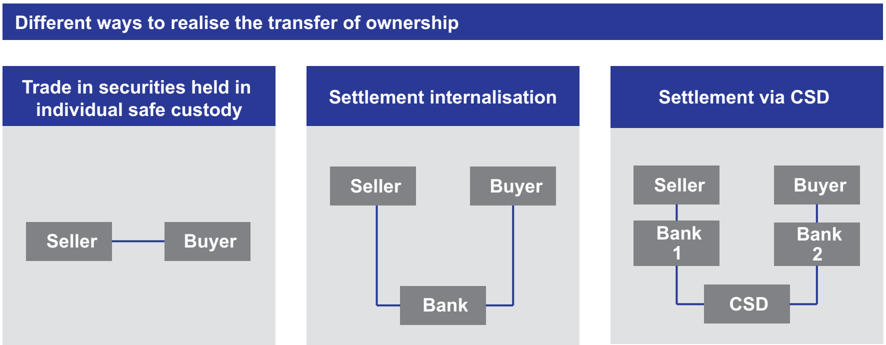


Fig. 6.12 Different models for effecting the transfer of ownership (Deutsche Boerse: The European Post-Trade Market, 2005, p.23. Available at: https://deutsche-boerse.com/blob/2534550/34b8a2d88a8b8e8bf6621fdf8513bc80/data/the-european-post-trade-market-0205_de.pdf)

and the final settlement of securities transfers. This situation is referred to as "settlement finality," a legally defined moment marking the point(s) in time at which transactions are irrevocable. Settlement finality, a key element in cases of bank insolvency or bankruptcy, is especially important in relation to the management of credit and liquidity risk.

Settlement activity can be measured by the number of transactions and/or the turnover amount (i.e., the total value of the settled trades). Figure 6.13 shows the number of transactions and total value of settled trades of CPMI countries by geographical region.

The European Central Bank (ECB), a thought leader in the field of settlement, makes use of three distinct definitions of settlement finality on settlement harmonization:

- *Settlement Finality I* is the moment of entry of a transfer order into the system, or the moment when a transfer order is protected against insolvency procedures.
- *Settlement Finality II* is the irrevocability of a transfer order (and not of the transfer itself) according to the rules of the system.
- *Settlement Finality III* is the moment of irrevocability of transfers (bookings in securities and cash accounts) according to the rules of the system.[8]

Harmonized standards for settlement finality are necessary to ensure efficient and safe cross-border trade and post-trade environments.

6.5.2 Types and Models of Settlement

Settlement is commonly executed according to one of the two alternative processes: firstly, by delivery free of (without) payment. This is a delivery of securities with no corresponding payment of funds. Two linked, free of payment instructions, can be referred to as "delivery versus delivery."

[8] European Central Bank. Harmonization. Available at https://www.ecb.europa.eu/paym/t2s/harmonisation/activities/html/index.en.html.

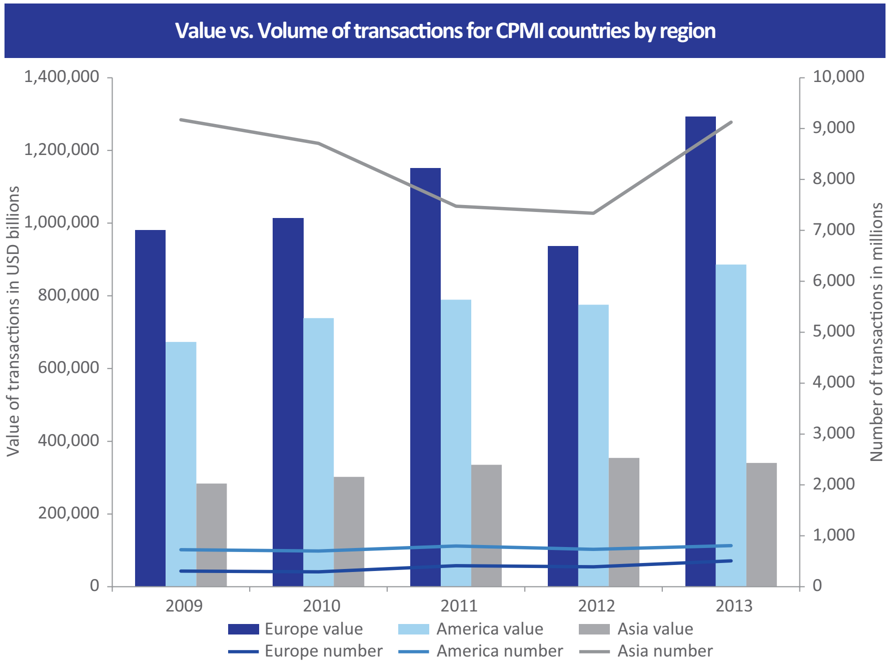


Fig. 6.13 Value and number of settlement transactions processed by CPMI CSDs by region (Bank for International Settlement, 2014. "Statistics on payment, clearing and settlement systems." Available at http://www.bis.org/cpmi/publ/d142.htm)

The second type is DVP, mentioned as a mitigant to credit risk in Sect. 6.3.1. DVP implies a link between a securities transfer system and a cash transfer system that ensures that delivery occurs if, and only if, payment occurs.

DVP usually takes the general form of a basic three-step process: First, the SSS blocks the underlying securities in the account of the seller, and then requests a transfer of funds from the buyer's bank to the seller's bank in the payment system (PS). Finally, it delivers the securities to the buyer if (and only if) a confirmation of settlement of the cash leg from the settlement bank is received (Fig. 6.14).

The CPMI report of the G-10 Central Banks, "Delivery Versus Payment in Securities Settlement Systems," published in 1992, identifies three approaches applied by SSSs to achieving DVP:

- *DVP model 1* are systems that settle transfers of both securities and funds on a gross (or obligation-by-obligation) basis. The final (irrevocable and unconditional) transfer of securities from the seller to the buyer occurs at the same time as the final transfer of funds from the buyer to the seller. The advantage is that transfers become final as they occur. That reduces exposures among users during the settlement day.

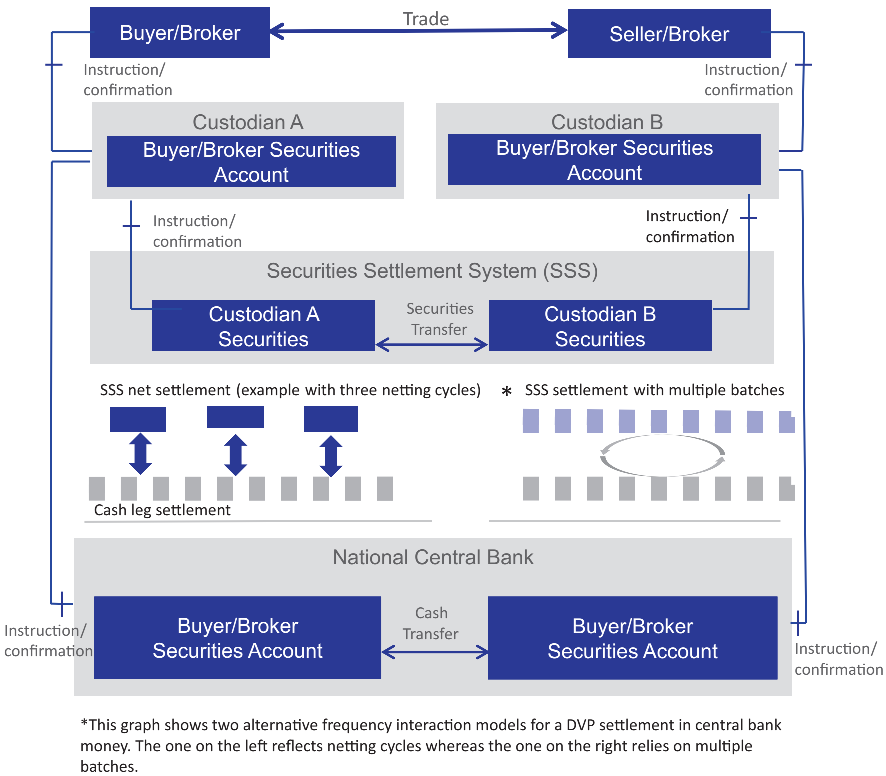


Fig. 6.14 The DVP process (Adapted from: European Central Bank: The use of bank money for settling securities transactions; ECB; 2004. Available at: https://www.ecb.europa.eu/pub/pdf/other/useofcbmoneyforssten.pdf)

- *DVP model 2* are systems that settle securities transfer obligations on a gross basis. The final transfer of securities from the seller to the buyer occurs throughout the processing cycle, but settles fund transfer obligations on a net basis. At the end of the processing cycle, the final transfer of funds from the buyer to the seller occurs. The advantage is that less cash liquidity is required as a result of the netting among users.
- *DVP model 3* are systems that settle transfer obligations for both securities and funds on a net basis. Final transfers of both securities and funds occur at the end of the processing cycle. The advantage is a reduction in the required cash and securities liquidity in contrast to models 1 and 2.[9]

The chart below shows the adoption of the different DVP models by geographical region (Fig. 6.15).

[9] Committee on Payments and Market Infrastructures, 1992. "Delivery versus Payment in Securities Settlement Systems." Available at http://www.bis.org/cpmi/publ/d06.pdf.

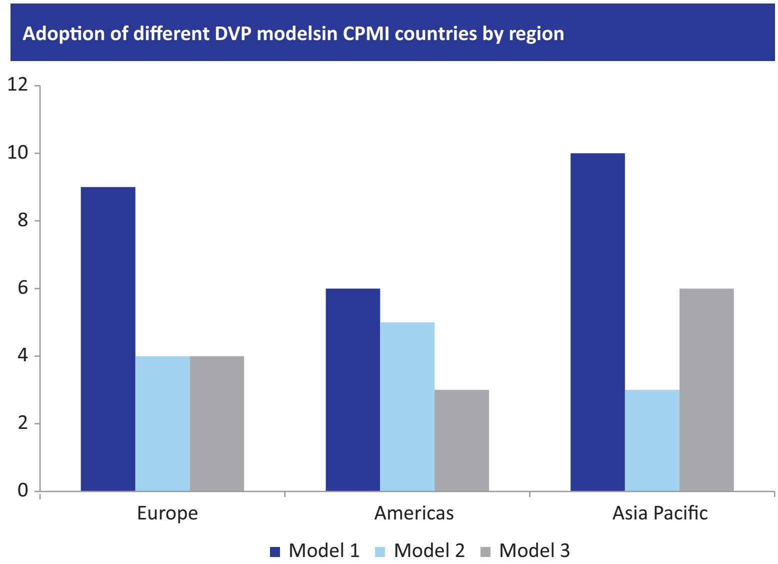


Fig. 6.15 Regional breakdown of DVP models, 2014 (Bank for International Settlement. Payment, clearing and settlement in various countries, September 2014. Available at: www.bis.org/cpmi/paysysinfo.htm)

6.5.3 *Choice between Two Qualities of Settlement Funds*

Simply put, DVP, as the term implies, is the transfer of cash against securities. Two distinct kinds of cash or money can be used. Firstly, there is central bank money, or money which is considered a liability of a central bank. Settlement in central bank money typically calls for the discharge of settlement obligations on the books of a central bank. Secondly, there is commercial bank money, or money which is considered a commercial bank liability, and is represented by the deposits held at the bank. Commercial bank money settlement carries a risk: settlement funds may not be available in the event of the insolvency of the commercial bank that is providing the settlement services. This risk is a function of the financial health of the commercial bank at stake.

6.5.4 *Cross-Border Settlement in Central Bank Money: TARGET2-Securities*

Launched in 2015, T2S is an ECB project to create a single European settlement platform. This platform is for the settlement of all securities executed in Euros, in central bank money, on a real-time gross settlement (DVP model 1) basis. The so-called integrated model will allow both securities accounts held by CSDs and dedicated cash accounts (DCAs) opened on the books of national central banks to be managed directly on the same T2S platform. The single platform is intended to overcome differences in national rules and requirements and remove technical, legal, and fiscal barriers that today prevent efficient cross-border clearing and settlement in the EU. The participating markets represent almost all Eurozone settlement activity.[10]

We next consider the general mechanics of the platform: buyers transfer cash from their real-time gross settlement accounts (TARGET2 for the Euro) to dedicated cash accounts (DCAs) held with their national central bank.

Trades can then be settled DVP, with instructions being validated and matched on the T2S platform. Payment takes place via the DCA of the buyer, provided that there are sufficient funds in this account. The trade can settle if the securities are available.

At the end of the day, any cash still in the DCAs is swept back to the RTGS account. Several optimization tools are also in place to improve liquidity levels and enhance settlement efficiency.

Payments in T2S are, therefore, effected by the individual member's national central bank in central bank money, that is, provided that the buyer has sufficient cash or collateral deposited there. Payments and the related collateral movements can be effected during the TARGET2 working hours (07:00–18:00 CET). All central bank money operations of the European System of Central Banks (ESCB) that provide liquidity necessitate the deposit, by the counterparty of the operation, of adequate collateral value in the form of securities. The deposit is made at the respective central bank, via a national or an international clearing system (Fig. 6.16).

Thanks to the economies of scale from consolidating settlement volumes from many platforms onto a single platform, T2S aims to reduce settlement fees in Europe. In addition to cost savings, the dynamic effects of T2S are also expected to shorten the custody chain and to facilitate a greater mobility of collateral. T2S will, therefore, also offer market participants a number of new opportunities.[11] A 2014 study by Oliver Wyman that was commissioned by Clearstream outlined the T2S benefits that banks

[10] AS Eesti Väärtpaberikeskus (Estonia), Bank of Greece Securities Settlement System (BOGS), BNY Mellon CSD SA/NV (Belgium), Centrálny depozitár cenných papierov SR, a. s. (Slovakia), Clearstream Banking AG (Germany), Depozitarul Central S.A. (Romania), Euroclear Belgium, Euroclear Finland Oy, Euroclear France, Euroclear Nederland, Iberclear—BME Group (Spain), Interbolsa (Portugal), KDD—Centralna klirinško depotna družba, d.d. (Slovenia), Központi Elszámolóház és Értéktár Zrt.—KELER (Hungary), Latvijas Centrālais Depozitārijs (Latvia), Lietuvos centrinis vertybinių popierių depozitoriumas (Lithuania), LuxCSD S.A. (Luxembourg), Malta Stock Exchange, Monte Titoli S.p.A. (Italy), National Bank of Belgium Securities Settlement System (NBB-SSS). Oesterreichische Kontrollbank Aktiengesellschaft (Austria), SIX SIS Ltd. (Switzerland). VP Lux S.a.r.l. (Luxembourg), VP Securities A/S (Denmark).

[11] PricewaterhouseCoopers AG Wirtschaftsprüfungsgesellschaft, 2013. "The 300-billion-euro Question: Survey on the Benefits of TARGET2-Securities." Available at http://www.clearstream.com/blob/6220/fea603b397e51f16a0256b31fda02ad2/migrated-9b3hc6580nsgden-t2s-pwc-paper-pdf-data.pdf.

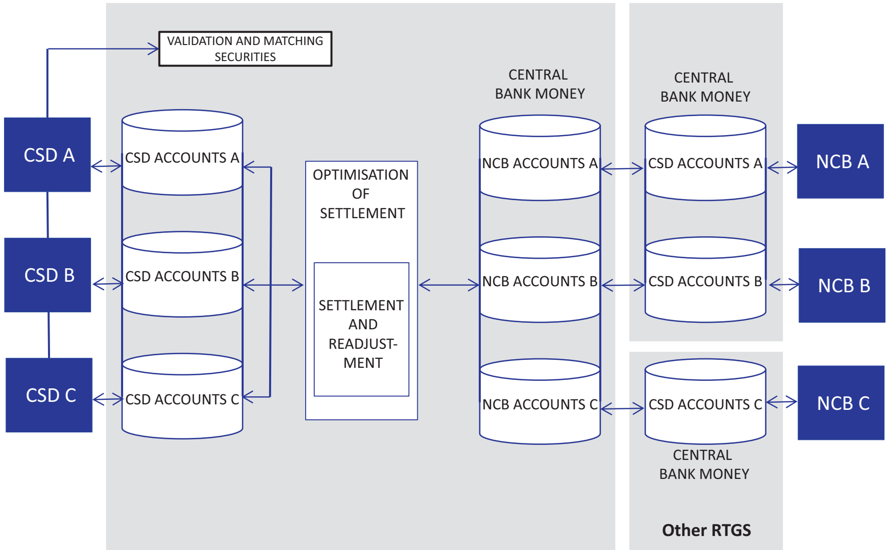


Fig. 6.16 The T2S settlement model (Banque de France. TARGET2-Securities. Available at www.banque-france.fr/en/financial-stability/payment-systems-and-market-infrastructure/target2-securities.html)

can unlock by consolidating their securities and cash holdings in Europe directly in CSDs and central banks.[12] According to the study, by so doing, banks are able to:

- Remove settlement-related exposures by interacting directly with market infrastructures and central banks, and also benefit from the self-collateralization of transactions and so reduce their credit needs.
- Pool collateral for settlement purposes and tri-party repo transactions (see Sect. 6.6.3 for further details) to reduce collateral buffers currently fragmented across markets.
- Net more cash settlements by using fewer central bank cash accounts to fund activities across markets.
- Simplify operations by leveraging a single CSD to access T2S markets.

Case studies reveal that banks can realize significant capital, funding, and operating cost savings thanks to direct market access and asset consolidation. The study estimated the savings potential in three high-level case studies, based on conservative assumptions.

In addition to cost efficiencies, a more consolidated T2S model can provide further benefits to banks, increasing stability and reliability of securities service

[12] Oliver Wyman, 2014. The T2S Opportunity: “Unlocking the hidden benefits of TARGET2-Securities.” Available at http://www.clearstream.com/blob/68228/9f9261051598b77e44bddf291d655859/t2opportunity-pdf-data.pdf.

operations, and reducing operational complexity and risks. To take full advantage of these benefits, banks need to fundamentally reconsider and alter their current operating models, especially with settlement, in the securities service area.

In addition to T2S, the CSD Regulation[13] introduces a T + 2 EU settlement cycle. This means that the settlement period will be harmonized, set at a maximum of 2 days after the trading day for securities listed on stock exchanges, or other regulated markets. Market participants that fail to deliver their securities on the agreed settlement date will be subject to penalties, and will have to buy those securities in the market and deliver them to their counterparties. Europe, as a region, has therefore taken the lead globally in making settlement a standardized process with significant market benefits.

6.6 Securities Financing

6.6.1 Regulatory Momentum

In the aftermath of the financial crisis, regulators rushed to strengthen rules and regulations to take risk out of the financial market and to strengthen banks. The Basel III and Dodd–Frank regulations require banks to increase their equity levels to improve their solvency in the event of a crisis. Accordingly, these regulatory requirements result in an increased demand for so-called high-quality liquid assets (HQLA). Access to these securities, possibly via a third party, has become a key securities service. Securities financing is a key tool to fully utilize increasingly scarce HQLA.

6.6.2 Basics of Securities Financing

Securities financing is the ability to borrow or to lend cash or securities against collateral. In securities financing, collateral comprises assets given as a guarantee by a borrower to secure a securities loan, and it is subject to seizure in the event of default. Collateral management is the handling of all tasks related to the monitoring of collateral posted by a borrower to meet an exposure (optimization, substitution, settlement instruction, reporting, processing of margin calls and returns, notification of corporate events, etc.).

[13] EUR-Lex. "Regulation (EU) No 909/2014 of the European Parliament and of the Council of 23 July 2014 on improving securities settlement in the European Union and on central securities depositories and amending Directives 98/26/EC and 2014/65/EU and Regulation (EU) No 236/2012 Text with EEA relevance." Available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.257.01.0001.01.ENG.

Securities financing is generally used to enhance yield, enhance settlement, or access liquidity. Eligibility criteria will vary, but collateral typically includes cash, bonds, shares, and certain mutual funds. Collateral management is a securities service that combines analyses of collateral needs and availability and settlement operational capability. The aim is to transfer the collateral to where it is needed. The best performing collateral management systems use algorithms to ensure the most efficient use of collateral (i.e., to use the cheapest acceptable collateral to meet an exposure).

Securities lending agreements and repurchase (repo) agreements are the main securities financing transactions. Both agreements resemble collateralized loans. Still, under bankruptcy law, their treatment is more favorable to collateral takers who can simply sell the collateral and avoid delays. Repo and securities lending agreements contain key information, including the size of the transaction, the interest rate, the type of eligible collateral, the haircut, the maturity date, and the counterparties. The haircut is of particular interest in the context of collateral management. The haircut corresponds to the difference between the value of the collateral and the value of the cash. For example, €100 of securities as collateral for €96 in cash means a 4 % haircut. The magnitude of a haircut is mainly a function of the quality and liquidity of the collateral.

The financial intermediaries that participate in repo and securities lending transactions can be divided in two groups: (a) custodians and (I)CSDs who act as securities service providers for the repo and securities lending markets and (b) the securities dealers. The second group are customers of the first. The focus here, consistent with this chapter, will be the services by the custodians and (I)CSDs.

6.6.3 Securities Services supporting Securities Financing

Firstly, let us look at services supporting the repo markets. It is helpful to distinguish between bilateral and tri-party repos. Bilateral repos are repurchase agreements between two institutions, usually with DVP settlement. The cash giver may access a custodian or (I)CSD to receive, track, value, and account for the securities. In a tri-party repo transaction, a third party—the tri-party agent—provides a suite of collateral management and settlement services. These include settling the repos on its book, valuing the collateral (haircut), and ensuring that the lender's collateral eligibility criteria are satisfied.

Settlement occurs in the books of the tri-party agent, who performs the collateral management. Bilateral repos are mostly used to obtain specific securities and to raise cash against these securities. Tri-party is more suited to general collateral transactions. In the USA, the role of the tri-party agent is assumed by JPMorgan Chase and the Bank of New York Mellon. Outside the USA, the role of tri-party agent is assumed by the ICSDs.[14]

[14] The European Repo council states in its 2014 biannual report that more than 90 % of repo in the USA are tri-party, versus slightly more than 10 % in the EU.

Cash givers with a tri-party repo agreement have this cash upside: instead of using their cash, they can reuse the collateral they receive from the cash taker as collateral for OTC-derivative exposures they have with other counterparties. This benefit is important because newly enacted legislation (such as EMIR in the EU) has specific collateral-exchange requirements for OTC derivatives not cleared by a central counterparty. Cash givers could also reuse the collateral to undertake a repo to obtain cash financing, or to support liquidity or treasury lines.

Secondly, an institution may also want to borrow a security to avoid "failing" on a settlement delivery. Fails lending brings significant market benefits. It increases settlement efficiency by manually or automatically lending securities to enable a settlement where insufficient securities were available for transfer to the seller's account. The securities are borrowed against collateral, i.e., cash or securities in the account of the seller. Under the fails lending programs offered by (I)CSDs, the borrower is charged a fee split between the (I)CSD and the lender, that is, typically another customer of the (I)CSD. Securities lending can also serve more strategic objectives and investment strategies. The process is similar.

Thirdly, as mentioned in the settlement section, some securities service providers, including the ICSDs, facilitate access to central bank money operations. Customers can use their eligible assets as part of their comprehensive services for accessing EUR and USD central bank liquidity, via collateral pledges, to the relevant central banks. Customers can use this service for central bank discount window borrowing, and to participate in tender offers and auctions. The ICSD acts as a neutral tri-party agent throughout the collateral management life cycle, all the way from instruction matching to collateral allocation, valuation, and substitution. The principal relationship remains between the central bank and the borrower.

The bottom line is cost savings. The combined effect of using securities as collateral in repo transactions as well as the lending of securities results in reduced overhead and, therefore, increased income for the market participant. The securities are being made to work instead of sitting inactive in accounts at CSDs and other custodians. This enables the market participant to make the most of their assets.

Indeed, lending and borrowing securities enhances market liquidity and settlement efficiency. Many anticipate that the increasing demand for the collateralization of exposures will lead to a relative scarcity of HQLA. Therefore, collateral management service providers are developing collateral management solutions their customers can use to mobilize their collateral across markets. A group of CSDs worldwide have formed the Liquidity Alliance to advance common solutions to the challenge of global collateral. (See box for further information.)

6.7 Custody Services

Custody services include safekeeping and asset servicing, or, as this is sometimes called, corporate action processing. These are traditional securities services that have evolved in complexity, and that remain essential to a secure and efficient

Liquidity Alliance
The Liquidity Alliance was established in January 2013 as a platform for CSDs to collaborate on collateral management. This group of CSDs offers members an opportunity to discuss key developments, identify business opportunities in collateral management, and share individual market experience. At the same time, the Liquidity Alliance promotes studies and industry research. Liquidity Alliance members are from different regions of the world, a fact that brings together a unique pool of global insight and expertise.

securities market. Custody services are also a critical part of any integrated securities service. Without custody services, the other securities services are of limited value; once issued or settled, securities must be held in a custody account and serviced. Similarly, efficient collateral management services are impossible without the ability to service the assets used as collateral.

6.7.1 Safekeeping

The majority of securities these days are immobilized within the (I)CSDs and are, in fact, dematerialized. Put differently, they are no longer represented by physical certificates, but instead by data entered into the systems of these (I)CSDs.

To be sure, it would be logical to conclude from these two facts that the need for the services of custodians is greatly reduced. However, this notion would overlook some of the key features of the securities service landscape. Some reasons why custodians are still used include:

Ineligibility: some market participants may be unable to hold an account directly with a CSD, because they do not fulfill the CSD's account-opening requirements. Many CSDs offer standard services to a limited number of locally based financial institutions. Requirements could also be based on operational capability—with some market participants unable to invest in the technological solutions to connect directly with the CSD.

Expertise and economies of scale: custodians, by holding the securities of a number of investors, are able to leverage economies of scale. Moreover, specialized custodians will also have expert working knowledge of the CSD and local market practices, a scenario which may be advantageous to market participants.

Specialized services: custodians often provide additional value-added services related to the custody of securities. These services can include additional reporting for a certain group of market participants.

By choosing to hold their securities via a custodian, market participants are taking a significant step: they are choosing to outsource asset servicing activities to an entity that can complete these tasks better—and cheaper—than the market partici-

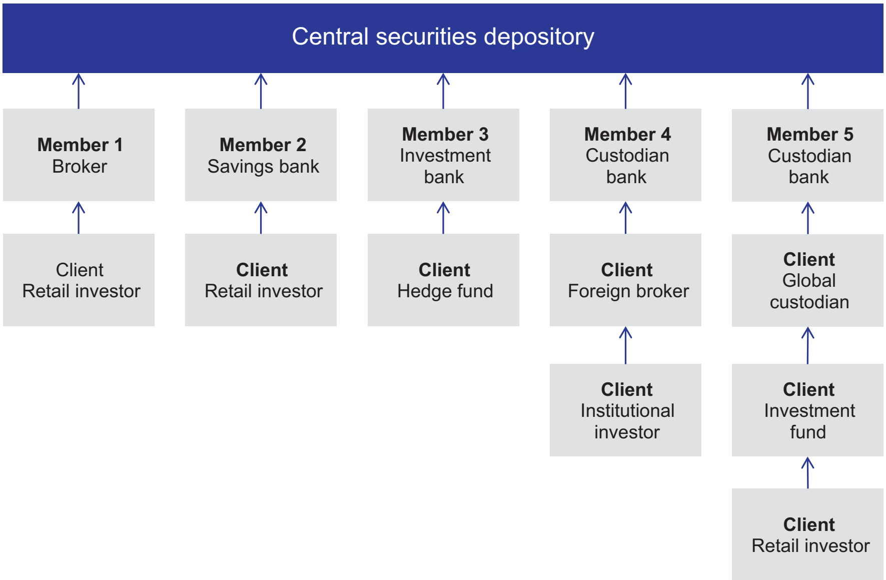


Fig. 6.17 Examples of multitiered intermediation in securities custody (Chan, D et. al.: The Securities Custody Industry, ECB Occasional Paper Series No 68, 2007, p.11. Available at: https://www.ecb.europa.eu/pub/pdf/scpops/ecbocp68.pdf)

pants themselves are able to do. Moreover, this decision on whether or not to keep asset servicing activities "in-house" results in multitiered intermediation in securities custody (Fig. 6.17).

In each tier, a choice is made on the account structure—whether to hold the fungible assets of different clients together without separating out ownership, or to hold the assets of individual clients in individual segregated accounts. Generally speaking, we can define three separate models for holding securities at the local market level; they are detailed in Fig. 6.18.

Lower levels of segregation reduce transparency, which makes it more difficult to identify the beneficial owner of securities. On the other hand, higher levels of pooling in less segregated accounts can offer significant economies of scale by netting and the aggregate processing of corporate actions. The importance of the economic efficiency of nonsegregated account types is demonstrated by the continued high levels of omnibus account structures around the world. This is especially the case for larger, more sophisticated markets like the USA, Germany, and the UK (Fig. 6.19).

The 2008 financial crisis ushered in an increased regulatory focus on greater levels of transparency in account structure, particularly in the USA and Europe. In an earlier 2004 paper on client identification, the International Organization of Securities Commissions (IOSCO), did not require that custodians examine the

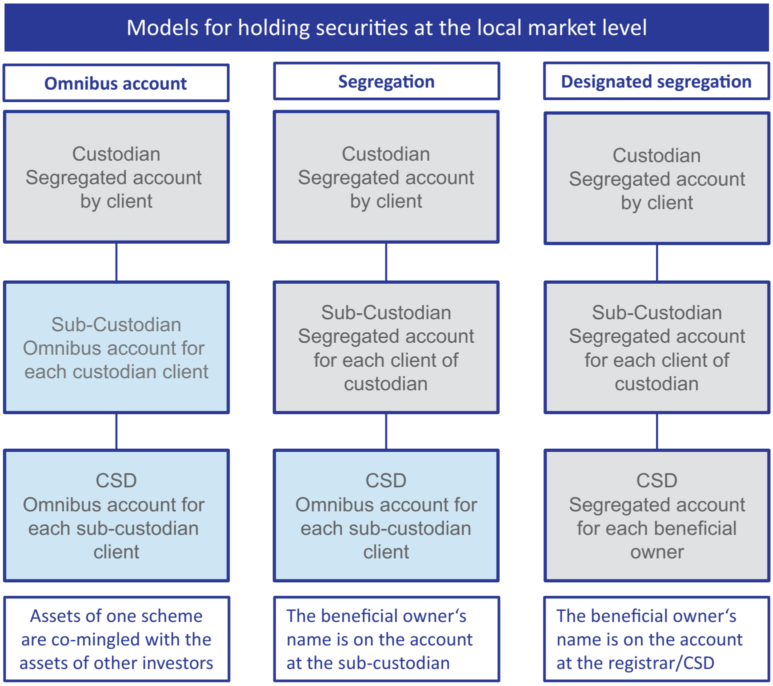


Fig. 6.18 Models for holding securities in the local market (Thomas Murray Data Services: CMI in Focus: Asset Segregation in CSDs, 2013. Available at: http://ds.thomasmurray.com/opinion/cmi-focus-asset-segregation-csds)

owners behind omnibus accounts.[15] And, until 2009, official guidance from the Financial Crimes Enforcement Network (FinCEN, a department of the US Treasury) stated that financial intermediaries were not required to look beyond their immediate counterparties either.

The balance between economically efficient account structures on the one side and transparent account structures on the other is likely to be a key regulatory topic in the coming years.

6.7.2 *Asset Servicing*

Asset servicing includes the handling of dividends for equity, and of income and redemptions for bonds, as well as the processing of corporate action events. In addition, various ancillary services are available, including withholding tax reporting services and proxy voting services.

[15] IOSCO: "Principles on client identification and beneficial ownership for the securities industry." Available at https://www.iosco.org/library/pubdocs/pdf/IOSCOPD167.pdf/.

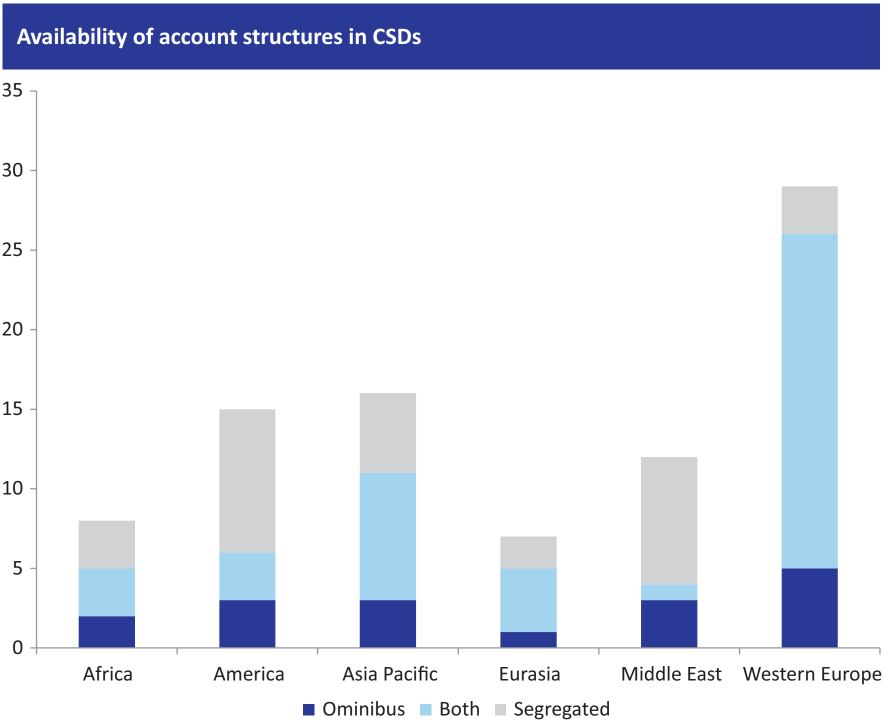


Fig. 6.19 Availability of account structures in CSDs (Thomas Murray Data Services, 2013. "CMI in Focus: Asset segregation in CSDs." Available at http://ds.thomasmurray.com/opinion/cmi-focus-asset-segregation-csds)

6.7.2.1 Income Payment

Interest paid to the bond holders is also referred to as income and coupon payments. Coupons are presented to the issuer for payment. Following the receipt of funds from the issuer, the proceeds are credited to holders' accounts on payment date, after deduction of applicable withholding taxes. Payments are usually made in the original payment currency as determined by the issuer.

Given that securities may be sold, or become part of a collateral management or securities lending transaction, the record date is important; it is the date on which the relevant system operator (e.g., the CSD) establishes which holders are recorded in the system as eligible to receive the coupon, or other entitlement, on a security. In the international market, the record date is usually the close of business, one business day before the payment date of the coupon (or other entitlement). For domestic securities, the record date varies for different security types according to domestic market practice. After the record date, securities movements are processed ex-coupon or ex-dividend.

6.7.2.2 Redemption Payments

Bonds are issued for a finite duration, unlike equities which have no predefined maturity date. The termination of a bond—that is, when it ceases to exist and the final payment is made by the issuer to the investor—is referred to as redemption. Redemption can be either total or partial. A total redemption can happen earlier than the final date, but a total redemption is a straightforward process—payment is made to holders provided that the issuer is not in default. Partial redemptions are slightly more complex as they require a level of "fairness" among holders. Some of the methods used to achieve this are described below.

In the drawing process, an algorithm is typically applied to distribute the total amount to be drawn from each account participating. In a redemption on nominal value, an equal part of all notes of a security is redeemed, the denominations being reduced accordingly. In a partial redemption with a pool factor, an equal part of all notes of a security is redeemed but the initial face value is not reduced accordingly. A ratio (the "pool factor") is assigned to the security, reflecting the face value of principal still to be redeemed. For each interest payment, the amount of interest payable is then calculated on the basis of the outstanding amount of principal, not on the basis of the denomination of the security.

Redemption proceeds provide important funds for settlement. Prompt payment is, therefore, crucial and is sometimes anticipated at the various levels of the custody chain. Intermediaries may, however, depending on the creditworthiness of the issuer, make payment conditional on receipt of funds from the issuer.

6.7.2.3 Corporate Action Processing

A corporate action refers to the processing of any event that impacts the rights of a company, its shareholders or bondholders, excluding income events like interest or dividend payments. It may be initiated by the issuer, a third party, or holders. For some corporate action events, holders must respond by selecting from a list of possible actions.

Corporate events can be divided into two broad families:

- Predictable events: events foreseen in the security's documentation (such as the terms and conditions), including wording around the event timing and deadlines. Examples of predictable events are conversion options, put options, or warrant exercises.
- Unpredictable events: events not foreseen in the security's documentation. They are announced and described in additional documents by the issuing company's management. Examples of unpredictable events include repurchase offers or stock splits.

Both predictable and unpredictable events can be subdivided into three main categories. These categories are based on whether the holder of the security has to take action on the event:

- Mandatory events: participation and consequences are compulsory, applicable to the whole outstanding amount held. No instruction from the security holder is required. Examples of mandatory events are stock splits, rights distributions, and mergers.
- Mandatory events with choice: participation and consequences are compulsory and apply to the whole outstanding amount held. However, a choice or option is available to the security holder. An instruction is only required if the security holder does not want the default option applied. Examples of a mandatory event with choice include mergers with choice or non-automatic bonuses.
- Voluntary events: participation and consequences are at the holder's discretion. An instruction is required if the security holder wishes to participate. The issuer will usually inform all holders of the event about to occur. Sometimes this notice is provided in the original offering documentation for the security. If no action is taken by the holder, the default action, as stated in the notification, is applied. Purchase offers, conversion options, or subscription offers are examples of voluntary events.

Some complex corporate actions may involve mandatory events tied with subsequent voluntary events (two leg events), for example, a mandatory rights distribution followed by a subsequent voluntary subscription offer. The entitled holding is fixed on the record date, or on the actual date, according to the terms and conditions of the relevant corporate action. A non-exhaustive list of corporate events is included in the annex to this chapter.

6.7.3 Straight Through Processing and Automation

Corporate action information must be collected and disseminated to relevant participants before it can be processed. This is relatively simple for mandatory corporate actions. Normally, it only requires that participants are informed of the event and notified that the corporate action is processed. The process is more complex for voluntary corporate actions that require choices from investors down the holding chain. Here, the level of automation in the communication between the involved parties becomes critical for reducing the risk of human error, and in increasing the speed and efficiency of corporate action processing.

Markets utilize different communication media to transfer information related to corporate action events. These media require different levels of manual processing. Channels requiring more manual input include, for example, fax, e-mails, or even conventional mail. There are also fully electronic systems that are mostly automatic. They are, therefore, capable of handling straight through processing (STP), or the elimination of any manual intervention between an event announcement and the action taken.

In many markets, participants now mostly communicate via SWIFT messages. SWIFT, the acronym for Society for Worldwide Interbank Financial

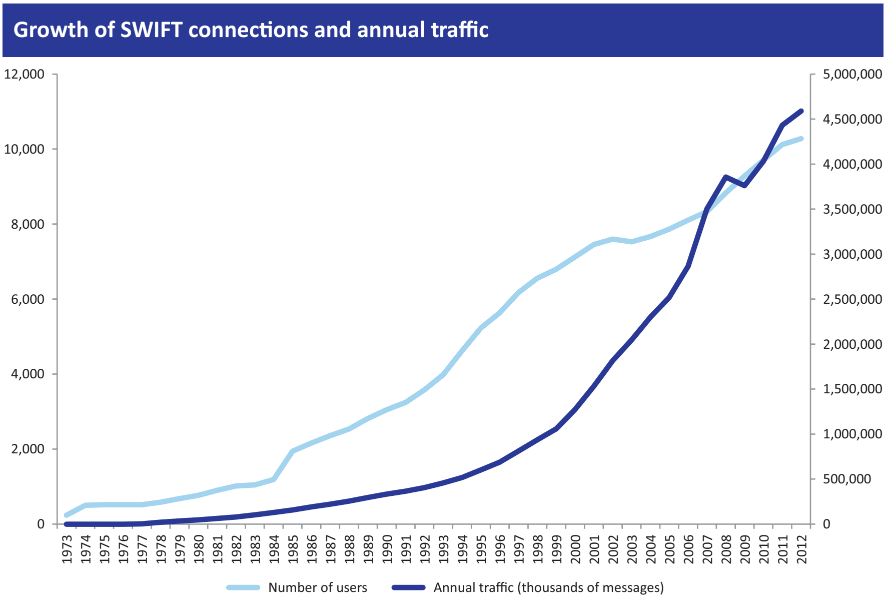


Fig. 6.20 Growth of SWIFT connections and annual traffic (Society for Worldwide Interbank Financial Telecommunication. SWIFT History. Available at https://www.swift.com/about-us/history)

Telecommunication, was initially established by banks to manage the secure transmission of payments internationally (Fig. 6.20).

SWIFT is the most widely used network for exchanging electronic financial messages. In 2014, more than 10,500 financial institutions and corporations in 215 countries were using the network. SWIFT enables its customers to automate and standardize financial transactions, a process that lowers costs, reduces operational risk, and eliminates inefficiencies from their operations. Transactions include payment, securities, and treasury activities.

SWIFT essentially provides an electronic, worldwide messaging service which enables financial institutions to exchange data quickly, reliably, and securely. The use of standardized messages enables financial institutions to automate their data processing. Many financial institutions and CSDs also develop their customer connectivity so that it is SWIFT and ISO compatible. SWIFT develops and maintains formats (e.g., ISO 15022 or 20022) that are strictly followed by financial institutions to ensure compatibility and interoperability. These formats are also the basis for the global, industry-owned association Securities Market Practice Group's market practice guidelines for how the messages are to be used globally in a harmonized manner.

The SWIFT network and standardization efforts represent a significant contribution to the harmonization and safety of financial communications.

Risks in the processing of corporate actions can be very significant. They can result in huge losses when there are errors in any of the links in the chain. Yet, so far, there has been relatively little progress in developing international standards for corporate action processing. The notable exceptions are CPMI-IOSCO principles and the CSD Regulation in Europe.

STP greatly reduces the risk of error caused by the number of intermediaries in corporate actions. STP also has a distinct advantage in ensuring that the complexity of the corporate events is handled in an efficient manner. Not surprisingly, new developments such as the SWIFT ISO 20022 format aim to reduce the amount of manual processing required to an absolute minimum.

6.8 Conclusion

In his speech before the European Parliament on 15 July 2014, Jean-Claude Juncker, the then candidate for President of the European Commission, said: "I believe we should complement the new European rules for banks with a Capital Markets Union. To improve the financing of our economy, we should further develop and integrate capital markets. This would cut the cost of raising capital, notably for SMEs, and would help reduce our very high dependence on bank funding."[16]

The integration of capital markets, on a European or global level, has many dimensions. Of top importance, harmonization of market rules and standards fuels market integration.

This chapter has covered securities services with a view to demonstrating their importance to safe and efficient capital markets. Removing friction from settlement or custody services is a key part of this agenda. This means the ability to buy, hold, and sell securities without friction from cross-border settlement or custody services. This also means the ability to use collateral in one market to meet exposure in another. A number of market initiatives (public and private) are promoting the harmonization of settlement and collateral management services. However, the much more heterogeneous market practices applying to custody services remain somewhat overlooked.

This chapter has also provided a basic description of securities services and the related infrastructure. Efficient issuance, settlement, securities financing, and custody services are critical to reducing risks for markets and their participants. Harmonization of securities services is also a prerequisite for efficient markets. Efficient markets, rather than the most convenient markets, is where capital can best be invested, and they are the most likely to create growth.

In 2001, the Lamfalussy Committee issued a powerful statement on the benefits and challenges of an integrated securities market: "The EU has no divine right to the

[16] "A New Start for Europe: My Agenda for Jobs, Growth, Fairness and Democratic Change," Opening Statement in the European Parliament Plenary Session, 15 July 2014.

benefits of an integrated financial market. It has to capture those benefits by building an integrated European market—in many areas starting from a very low level."[17]

Europe illustrates the integration that needs to take place at a global level. In the years between the two above statements, progress has been slow. The last financial crisis of 2008 acted as a catalyst to risk reduction and standardization of market practices. Most are driven by regulation; some are driven by industry initiatives. Securities services, like the markets they serve, are truly in transition.

Acknowledgment The author would like to acknowledge the valuable contribution of Ms. Ann-Kathrin Treis, Ms. Brenda Nolden, and Mr. Christopher Hollifield and thank them for their continued support.

[17] Lamfalussy Committee, "Final Report of the Committee of Wise Men on the Regulation of European Securities Markets," 15 February 2001, p. 8.

Chapter 7
Information Technology

Matthias Kluber

7.1 Introduction

Superior *information technology (IT)* is the essential success factor in any exchange organization worldwide, regardless of the instruments being traded for a panoply of asset classes that can range from shares and bonds to derivatives and commodities.

The days of trading pits with brokers milling around in colorful jackets, taking client orders over the telephone, are history. A modern stock exchange today is first and foremost an IT service provider.

In this brave new world of advanced technologies, the following key characteristics will determine the service quality of a stock exchange. Together, they will drive the design of its underlying IT systems:

- *Reliability*: A stock exchange is a critical component of the macroeconomic infrastructure, comparable to transport systems, communication networks, and energy supply. Every day, millions of investors rely on the *availability* of equity markets, and on the predictable execution of their orders.
- *Transparency*: The exchange should provide complete and timely information to the market about the operational state of its systems and the market's behavior. The relevant information includes the status of the current order book, and of individual member transactions and traded prices.
- *Integrity*: The exchange technology must prevent unpredictable system behavior even in exceptional circumstances, such as uncontrolled process flow by automated trading programs (Mad Machines) of "member installations" or faulty orders (Fat Fingers). Orders that cascade in an uncontrolled way because of these exceptional circumstances may lead to brief bursts of extreme market activity and, in so doing, can trigger a *Flash Crash*.

M. Kluber (✉)
Deutsche Börse Group, Mergenthalerallee 61, Eschborn 65760, Germany
e-mail: Matthias.kluber@deutsche-boerse.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_7

- *Fairness*: All exchange members are treated equally. In today's markets, exchanges achieve fairness by transparently offering a menu of standardized connectivity options, rather than by having a one-size-fits-all interface with the exchange.
- *Low Latency*: As markets move at ever faster speed, members rely on immediate system response and instant transaction processing. This kind of transaction processing, provided at the highest speed enabled by the latest technology, is particularly relevant for *high-frequency trading (HFT)* and *algorithmic trading*.
- *Predictability*: Members expect consistent system performance irrespective of the system load. In fast market scenarios in particular, systems should operate as usual, i.e., without any delay in transaction processing and market data distribution. Performance may degrade under exceptional volumes in other systems, but the same does not hold for exchange systems because they need to be highly scalable and maintain sufficient headroom to cope with peak loads.
- *Easy Access*: A regulated public exchange should be open for a diverse set of trading members, each with different business models and investment motives. The connectivity options should correspond with various technical requirements and expertise as well as the members' geographical locations. Technical barriers should be minimized for market entry to fulfill this easy access, e.g., with so-called *Zero Footprint*[1] connections.

These key characteristics apply to a broad range of market models and exchange systems. The specific characteristics of each equity market and its membership structure will ultimately determine how these principles are applied by the IT systems. For example, in a traditional floor-trading environment, low latency would signify the prompt display of prices on a screen within a few seconds after a trade is executed. In today's high-performance trading systems, transactions are processed end to end in less than a thousandth of a second. A billionth of a second can matter hugely for some members who are deploying market-sensitive trading strategies and algorithms.

These aforementioned design principles must be manifested in the building blocks of the exchange's technical environment (see Fig. 7.1). We will take a closer look in the corresponding sections.

However, such design principles require substantial capital investment for their implementation, from concept to reality. In the process, they often even compete with each other as we will see in the following sections.

- Core processing is the heart of exchange functionality. This is where order books are maintained and trades are executed by matching orders according to the rules of a market model.
- Transaction interfaces and market data interfaces are both critical for secure and fair member access to the exchange functionality; they keep the architecture efficient and scalable. Standardized *gateways* manage the market members' access to the core processing.

[1] Zero Footprint connections do not require special exchange software or hardware installations and maintenance at the member site.

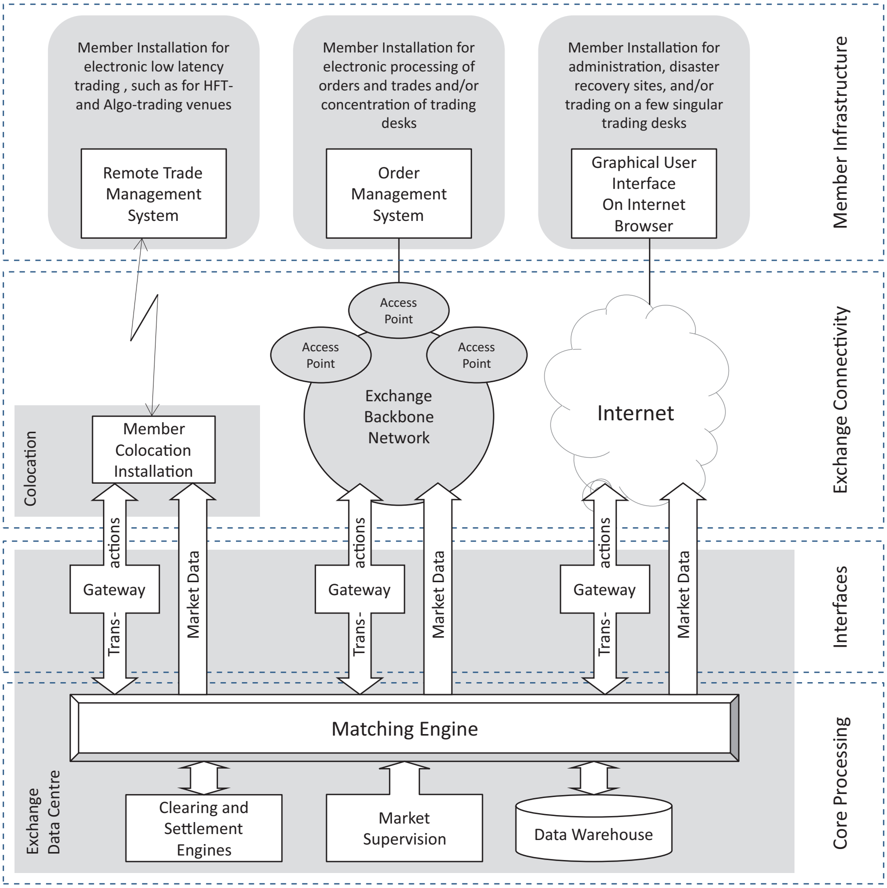


Fig. 7.1 High-level overview of building blocks in an exchange system

- Exchange connectivity is the connectivity options that member firms may deploy; they are compatible with the scale of their operations and trading and investment style. Exchanges typically offer a highly portable access option via the Internet that is suitable for "human traders" at smaller firms, or for use at *disaster recovery* locations. Larger trading desks and client-driven order routing businesses would generally connect via access points in an exchange's *Wide Area Network (WAN)*. Proprietary traders pursuing short-term strategies with high transaction throughput and extremely fast response requirements often opt for a *co-location* site. In a co-location facility, their trading engines that are controlled from a remote trade management installation reside in close physical proximity to an exchange's core processing center.
- Member infrastructure is the technical infrastructure that members have to build and maintain to connect to an exchange.

In the closing section of this chapter, Sect. 7.7, I describe how exchange organizations measure, control, and publish system performance information.

7.2 Core Processing

In the 1990s, the first generation of electronic exchange trading systems specifically designed for high availability made use of specialized *computer operating systems* designed for uninterrupted service with minimal downtime for maintenance. Many exchanges deployed Tandem NonStop[2] and OpenVMS[3], both of which are now part of the Hewlett-Packard Company. Today, state-of-the-art trading systems are typically built on Linux, the Unix-like computer operating system. Linux is "open source," meaning that access is based on a model of collaborative software development[4]. Red Hat or SUSE and other vendors select from the existing Linux modules and hardware drivers to package complete distributions that conform to their customers' needs and the available *server* hardware. Because the Linux software is free, vendors generate revenue mainly by offering software maintenance contracts. They provide support services and will deliver software patches in the case of software bugs, or to offset any incompatibilities between software modules and the hardware.

High-performance trading systems, unlike most general computing environments, are not built upon software *virtualization* layers. It should be noted that these layers would shield the application code from the underlying server hardware and the computing in the *central processing units (CPU)*. This virtualization is very useful for mainstream computing in optimizing hardware utilization, facilitating software development, minimizing maintenance efforts, and, therefore, reducing costs. However, because this adds overhead in the processing, in liquid exchange markets, virtualization is inadequate under the extreme performance requirements for a *matching engine*. High-performance trading systems are generally designed to operate at extreme speed, without delays, even under high load. Consequentially, the capacity specifications are laid out with ample headroom. The utilization of system resources should be on the low end to avoid bottlenecks at peak loads. Not surprisingly, exchanges use high-performance servers with multi-core processors; and interconnections between servers have high *bandwidth*, at least 10 Gigabit per second, or more.

The technical setup of the server hardware also has to be optimized: Regular system maintenance activities, from hardware memory checks to fan control, are technically controlled via so-called system management interrupts in the computer operating system. What are interrupts? These will temporarily stop application processing and, in so doing, allocate resources to these maintenance activities. By fine-tuning these interrupts, system performance becomes more predictable.

---

[2] Tandem NonStop was introduced in 1976 and includes a server line as well as the integrated computer operating system NonStop OS.

[3] The computer operating system OpenVMS's predecessor VAX/VMS was released by Digital Equipment Corporation in 1977.

[4] Open-source software is made available with a license in which the copyright holder allows to use and to change the software for free. Open-source software is often developed in a collaborative public manner.

Linux vendors typically assemble two types of software distributions: a general-use distribution that utilizes a stock *kernel* and *real-time* kernels deployed in high-performance environments that depend on extreme processing predictability: In this latter type, the process scheduler within a computer operating system allocates time slots to various processes. As long as the CPU is busy with a process, all other processes have to wait for their next time slot. Real-time kernels force interruptions of the active process on a very granular basis. The result is that all other processes frequently have the chance to become active, and to react to events. Consequently, this makes the reaction time more predictable overall. The downside is the additional overhead attributable to more regular switching between the different processes (Fig. 7.2).

The average processing time might be increased using real-time kernels, but the predictability of the processing time is optimized and the fat tail of the performance distribution is minimized. A fat tail refers to outlier events, e.g., a transaction taking exceptionally long to complete.

The processing of financial instruments with separate order books (e.g., for different shares) can be distributed on separate physical server hardware. Even if the processing occurs on one physical server, it still can be scheduled in parallel by different threads of instructions (one per order book) on multiprocessor systems. Trading systems can therefore be very scalable. The impact of hardware failures or performance issues can be contained within a subset of instruments.

Duplicating key components of the trading architecture is the way to maximize reliability and availability in modern exchange trading systems. Order books can be maintained in two instances, primary and backup, and the transactions processed in parallel on both instances. One component will then be actively used while the other runs in standby mode. This process allows for a seamless failover in case of a

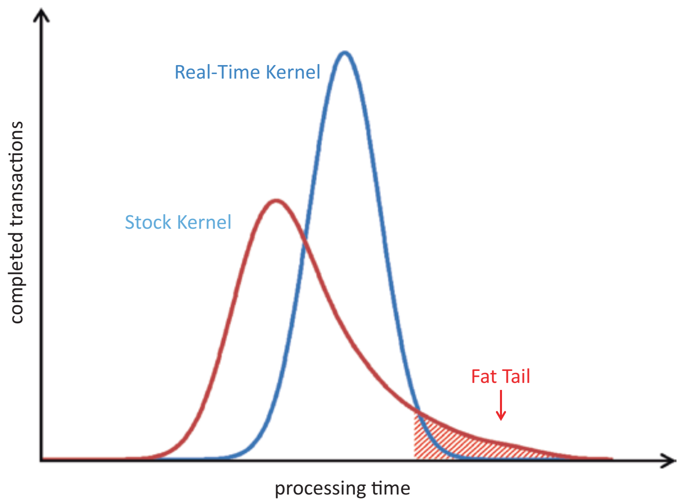


Fig. 7.2 Processing performance with real-time kernel versus stock kernel

defective hardware component. There is an upside to such *scalability* and reliability: High-end trading systems can just about host an unlimited number of instruments and asset classes, and support different market models, multiple market supervision entities, and diverging trading calendars.

However, there is an important limiting factor for the trading system capacity: All transactions for a specific order book[5] must be processed sequentially due to time prioritization of matching events during continuous trading. Therefore, the order book for any instrument must reside on one specific location in the memory. Any changes to the order book can only be consecutively applied, one after another. Distributing the processing of order book updates on more than one processing unit would require that these distributed units lock the memory containing the central order book for the time of the update, i.e., prevent other processing units writing to the memory. This is time consuming, even if measured in microseconds; it also limits the maximum throughput. Hence, state-of-the-art trading systems today concentrate the core matching for one instrument on a single CPU to avoid this extra time and expense. The corresponding order book information should reside in the Level 1 cache, i.e., the fastest memory closest to this CPU. The time to process an order book update by this CPU will then be the overarching limiting factor for a liquid instrument in the entire exchange system.

Pipelining: The concept that balances overall system throughput and the time span of individual order book transactions is called "pipelining." To optimize transaction times, one should ideally take all the steps in an order book transaction with a single CPU, and within the associated Level 1 cache. These steps include preparing the change in the order book—for example, receiving, decoding, and checking the transaction data—and then the update of the order book itself along with certain follow-up steps. The follow-up may consist of generating the relevant market information and the synchronous response to the member, encoding and sending the transaction data. Processing order book transactions end to end in this manner would block the CPU until all steps are completed. It is of interest, therefore, to identify some elementary steps that can be distributed over several CPUs within one matching engine. Breaking up transactions into a series of elementary steps is known as "pipelining." In this way, certain elementary steps can be distributed and executed in parallel by different CPUs, rather than by processing each entire transaction sequentially. More system overhead time is used in distributing these steps, because of the required data transfer between the CPUs. Nevertheless, each individual CPU uses less time than otherwise for the entire individual order book transaction.

Now suppose that the architecture of a high-performance matching engine can arrange an order book transaction into individual steps with a processing time of 15 μs (microseconds, a millionth of a second) for the longest single step (shaded step in Fig. 7.3). This order book can be updated approximately 67,000 times a second. The sum of all individual processing times (i.e., the processing time of the entire order book transaction, including the 5 μs overhead time per step) would be 75 μs. However, the additional system overhead times would be avoided if all steps

[5] The order book of a traded instrument is the list of the interests of buyers and sellers with price and quantity.

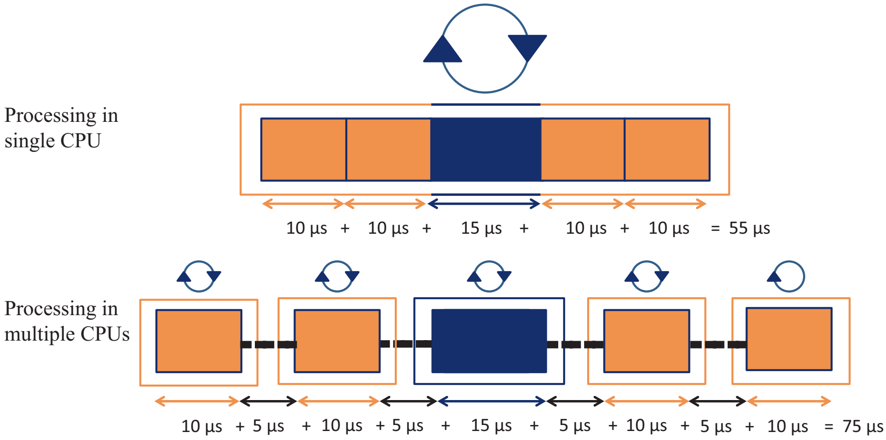


Fig. 7.3 Total processing time per order book transaction

were executed by a single CPU. That, in turn, would reduce the order book update transaction time to 55 µs. Nevertheless, in this case the overall system throughput would then be reduced to 18,000 transactions per second.

This trade-off between elementary processing speed and throughput capacity illustrates how exchange systems must be designed and optimized for specific use cases and market situations. In an exceptionally fast market, it would be conceivable to receive more than 18,000 order book update requests per second, or 18 order book update requests per 1 ms. On a single CPU processing under the assumptions that underlie Fig. 7.3, some transactions would have to be stored in *queues*, waiting for sequential processing, which would negatively impact the performance of the market. Here, an optimization for more throughput seems an appropriate response. If throughput is not as important as the reduction of transaction time, the trading system architecture should steer clear of cutting the transaction into such small pieces. In our example, a single order book update transaction, if processed on a single CPU in the core matching engine, can be accelerated from 75 to 55 µs, end to end.

Queue Handling: Generally, queue handling is a difficult challenge in the design of trading architecture. Typically, the sizing of a high-performance trading system will cater to ample headroom to avoid queuing and other capacity bottlenecks. Trading systems, even in fast market situations, should not slow down at all. However, in exceptional circumstances, the traffic flow might become congested. Sophisticated mechanisms need to allow the system to respond in an elastic manner. A minor system stutter may otherwise become amplified and eventually lead to an overall standstill. *Flow control* models (similar to models used for road traffic simulations) allow the trading system to gracefully slow down temporarily. Any performance degradation or slowdown represents an undesirable state for a trading system. But an escalating capacity overload and eventual standstill of the entire system are even worse. It must be unconditionally avoided.

To this end, customizable transaction limits are kept at the system gateways to prevent members from sending excessive transaction volumes. Transactions can be limited in two ways: The maximum rate of incoming transactions can be specified, or the number of open, not yet completed transactions per member can be limited. In both of these extremely rare cases, members will receive an error message from the gateway if they try to send a transaction that exceeds these limits.

We will now describe other mechanisms that can reduce system latency even further.

Optimistic Response: One of the most time-consuming aspects of transaction processing is writing information to a secure and persistent storage medium. An information update usually is synchronously stored on a storage disk, or other hardware device. To achieve an even higher confidence level, the data may then be copied to a second storage device in a geographically separate location. Once these written instructions are completed and confirmed, a transaction will be finalized and a response sent to the initiator. Finally, once this response is received, an initiator can rest assured that his or her transaction has been completed and safely stored.

To accelerate processing, a trading system can permit members to request an "optimistic response," as soon as the transaction is processed in the CPU, and once the order book is updated in the Level 1 cache memory. With this setup, one will receive a quick response; however, this response may not be reliable. The information in the memory could be irretrievably lost if, for example, there is a hardware problem. Alternatively, the member will receive the response once the order book update has been written reliably on a storage device. However, the storage device may also be lost if there's a larger disaster in the exchange data center. These responses, therefore, can only serve as a preliminary indication of the successful completion of the transaction. The legally binding confirmation of orders and trades will have to follow after the information is copied to the storage device in a second, geographically distinct data center. As an example, synchronous copying of data onto a fast, solid-state disk in a second data center 100 km away will take approximately an additional millisecond.

Transferring messages from one server to another is another source of latency in a trading system. For distributed computing in particular, multiple messages must be sent between clusters of servers. Not surprisingly, the standard communication protocols will add substantial overhead time to the transaction times. Once again, these are overheads measured in a few microseconds. Nonetheless, in a communication cluster with several nodes, these can result in a significant expansion in processing time.

Exchange operators therefore pay special attention to the messaging architecture for the transfer of data between processes. For instance, an incoming transaction related to a specific instrument must be forwarded to the matching engine that hosts this instrument's order book. Market data in turn must be sent out to the various member interfaces. The messaging architecture can either be customized for the exchange, or a low-latency vendor solution could be adopted. For trading systems with high throughput, it is essential to avoid a configuration with a central dispatch function that distributes incoming transactions to their target matching engines. A central broker in this approach would once more create a bottleneck. A distributed messaging middleware using *IP Multicast* technology and *remote direct memory access (RDMA)* leads to higher scalability and resilience. Transmission overhead can be significantly reduced by deploying RDMA. Moreover, a process in one server can write data directly into the memory of another server without involving their operating systems. This leads to high-throughput and low-latency networking, which is especially useful in massively parallel computer clusters.

Here's an example of the effective use of RDMA technology: The *InfiniBand* architecture of interconnecting computers with high-speed links and low latency to transmit data between each other via IP Multicast protocol.

Tuning these high-end trading systems for the highest possible performance and throughput described above effectively minimizes execution risk for exchange members. Nevertheless, further safeguards are required to prevent unintentional market movements such as Flash Crashes. In today's breathtaking speed of computer-based trading, human supervision of the market can be far too slow to control sudden and challenging market developments. The rare but much publicized Flash Crashes highlight how massive losses in market capitalization can occur within a blink of an eye. There are numerous possible causes for Flash Crashes: a programming error in an algorithmic trading engine (the "mad machine" phenomena), or an erroneous (fat finger) order entry by a *screen-based trader, to name* two.

Several safeguards for these risks are described in the following section:

- Transaction limits
- Functional checks
- Member-triggered emergency exits
- Function of *volatility interruptions*

If the number of transactions from an individual member exceeds predefined limits, a first line of defense against mad machines and fat fingers is the ability of gateways to reject transactions from this member, or even to disconnect the member's session. A second line is functional checks and predefined thresholds in the system. If a trader who is only authorized to buy or to sell up to a certain value accidentally confuses quantity and price, he might just not be able to send an order.

Sophisticated trading systems support the configuration of detailed authorization schemes, including risk and exposure limits for individual groups and specific traders. If certain limits are exceeded, these will first provide warnings, and then slow down or stop a member. Trading systems may also allow members to introduce price reasonability checks so that the entered price does not significantly differ from the price on the market.

Members must have control over their market exposure, particularly in exceptional situations. The emergency exits they need include stop buttons for clearing members (which will cut off some or all of the traders under their sponsorship), market maker protection parameters, and the automatic cancellation of orders if the technical connection to the exchange be interrupted.

The most important safeguard against Flash Crashes potentially is volatility interruptions: Here's how a volatility interruption works: If a market in an instrument moves so quickly that its price shifts outside a predefined range, a trading system can automatically halt the continuous trading. The market supervision team then has time to initiate an auction, so that members may review their positions and adjust pending orders before continuous trading in the instrument resumes.

7.3 Transaction Interfaces

Exchanges generally provide members with a range of different transaction interfaces. *Front-end trading applications*: Via these interfaces, members can submit transactions to the exchange trading engine, gateways being the standard entry points for the transactions. They also serve as a *firewall that* shields core processing from direct connections to the members.

Exchanges use various concepts for gateways. If a single gateway per traded instrument is configured (and all transactions from all members are directed to this sole gateway), the exchange can easily serialize all transactions for an instrument. In this way, it can implement a first-in, first-out service. For high availability, this single logical gateway would typically be implemented in a redundant hardware setup. However, if the exchange wants to support many members and instruments in a high-performance environment, connectivity via a single logical gateway does not scale well (Fig. 7.4).

When gateways serve many instruments, exchanges can evenly distribute member connections across the gateways as shown in Fig. 7.5. The gateways intermediate member connectivity, and relieve the core matching engine from supporting many individual member connections. That is because many members connect to a single gateway, and the gateways in turn connect to the matching engines.

Even with standard gateway hardware, the concept in Fig. 7.5 may still lead to slight variations in the gateways' and associated network links' performance. Time priority is assigned to orders only when they reach the matching engine. Therefore, extreme latency-sensitive members will always try to identify the "fastest" gateway at any point in time, i.e., the gateway through which they can reach the matching engine first. These members may choose to establish sessions on all of these gateways and then send their transactions in parallel, or to use their own methods to identify the fastest gateways, such as analyzing technical performance data.

Minimizing the impact of technical requirements by the exchange on member infrastructure and architecture is one general design principle of interfaces. Traditionally, exchanges have required members to install and maintain specialized exchange hardware or software (or both) on their premises. Today, members no longer need to install exchange hardware or software to connect to the exchange's *back end*. This is called "Zero Footprint Access." Modern trading architectures can be accessed without the need for specific hardware, operating systems, programming language, and compiler versions. That's as long as they support the general communication components, like *TCP/IP*.

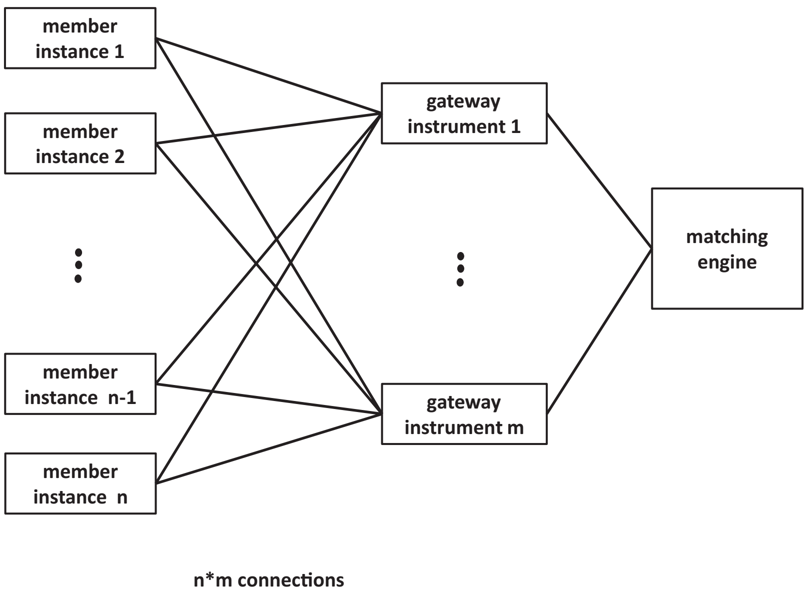


Fig. 7.4 Single dedicated gateway per traded instrument

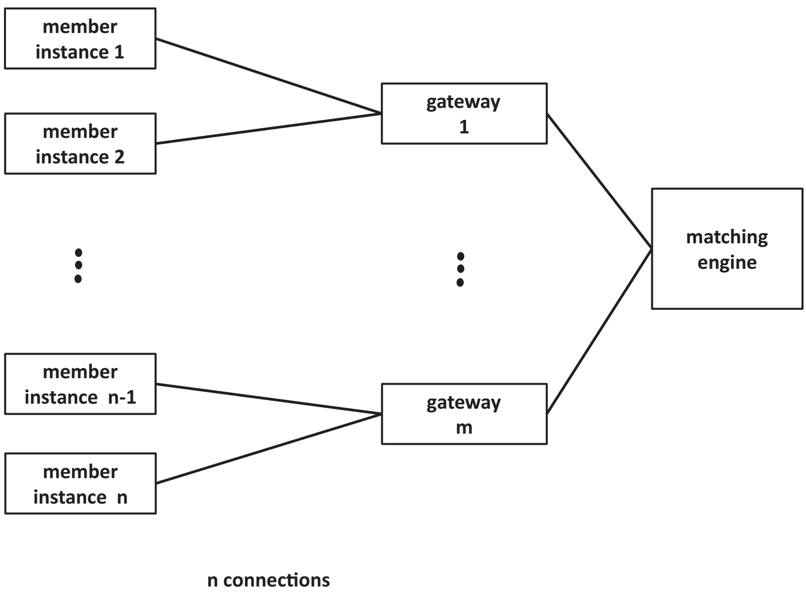


Fig. 7.5 Member connections to dedicated gateways serving multiple instruments

Transaction interfaces on exchange systems are typically asynchronous, message based, and session oriented. Members order their sessions from the exchange. Member software applications are connected to the trading system by opening a TCP/IP connection to an exchange gateway.

There are different design approaches for these interfaces and the corresponding gateways: Exchange proprietary interfaces allow high-performance access and full trading functionality. They support market making/quoting and additional services such as trading support information, or member-specific risk control messages. These proprietary interfaces are for members who require the highest throughput and the lowest latency. Messages exchanged between the member and the exchange across proprietary interfaces are, nonetheless, very similar to the standardized layout and content definitions of the *Financial Information eXchange (FIX)* protocol. The FIX protocol is optimized for traditional buy-side investors rather than for proprietary traders or market makers. Therefore, messaging efficiency can often be enhanced by small deviations from the FIX standard. These customizations may include a proprietary session layer with modified message headers, trailers, or additional user-defined fields and messages. The result is improved efficiency and performance that allows functional gaps in the protocol to be filled.

Exchanges may further support various session types within their proprietary interface specifications, for example:

- High-frequency sessions
- Low-frequency sessions
- FIX sessions

These session types can differ by their throughput limits and functionality. The pricing of these session types may reflect the way a member makes use of the exchange's technical infrastructure.

Members might submit a large quantity of order messages and other transactions to the trading system, resulting in a relatively small number of trades. The ratio of system transactions to trades will often exceed 100. Indeed, an exchange may also charge members for the number of transactions they are allowed to submit on this premise: The required capacity and the cost of the trading system depend more on the message volumes and less on the number of actual trades executed.

The high-frequency sessions offered by some exchanges are intended for market makers and HFT firms. These sessions accept higher transaction rates and allow members to make more intensive use of the exchange infrastructure. To minimize latency, the corresponding high-frequency gateways will, for example, accept only non-persisting orders, i.e., orders that are only kept in the Level 1 cache memory and not synchronously written to a storage disk. Data replication and recovery of trade events are restricted. The hardware of these gateways consists of powerful, dedicated, stand-alone servers that support special features like real-time kernel (see Sect. 7.2), kernel bypass, and *field-programmable gate arrays (FPGA)* for optimized latency and minimized variance.

A kernel bypass (a mechanism on network interface cards) allows data packages to be transferred straight to the application without being buffered in the operating system's memory. FPGAs allow configuring and optimizing microchips for very specific use.

Low-frequency sessions allow more functionality but, at the same time, they also restrict the number of transactions a member can submit. In addition, some exchanges offer special back-office sessions that serve only a subset of the low-frequency session functionality (mainly trade confirmations). The server hardware for the corresponding gateways will be less rigorously optimized for latency and performance.

Exchanges may also offer access via FIX gateways as an alternative to proprietary transaction interfaces. Members may prefer a FIX connection in order to standardize their connections to various exchanges. This is a point-to-point service based on the technology and industry standards of TCP/IP, FIX, and the FIX session protocol. The FIX protocol is not as flexible and efficient as an exchange proprietary interface, and it may limit performance and functionality. For example, a standard FIX session will not support the full scope of functionality for market making and quote submission that most exchanges offer. The exchange might offer two kinds of FIX sessions depending on the intended use of the FIX interface: (1) for order management and (2) back-office FIX sessions for the receipt of detailed trade confirmations organized by member business units.

7.4 Market Data Interfaces

There is a fundamental component for a fair and reliable market: An exchange system must provide order book and trade information as rapidly and transparently as practical to members. Order book information will be made available up to a specified depth based on the member's requirements. The order book data may either be refreshed upon each single order book change, or else be sent via a consolidated update that transmits all order book changes within a certain time interval. The consolidated update method can save bandwidth and be used for highly liquid order books.

Most exchanges use IP Multicast to broadcast market data given that all members should receive the same data simultaneously. IP Multicast is a method of sending data packages to a group of intended recipients in a single transmission. These packages are automatically copied within the network and distributed to several destinations based upon a receiver's subscriptions, in contrast to the TCP/IP protocol for individual transmissions between one sender and one receiver.

In trading systems, members subscribe to the market data streams for certain groups of instruments. However, IP Multicast packages are not necessarily delivered in sequence and lost packages are not automatically resent. That's because they are transmitted via the unreliable *User Datagram Protocol (UDP)*. IP Multicast transmission may generally work predictably and without interruption, but there is no flow control mechanism that guarantees delivery of a package. In fact, the receiving system at the member site will have to observe the sequence number provided by the exchange system and identify potential gaps, or correct the sequence of the incoming data stream. An exchange system, seeking to cope with the potential loss of IP Multicast packages, will typically disseminate its market data from the matching engine via two distinct IP Multicast streams over two separate network connections. A member system will then listen to both streams, and forward the IP Multicast package which it receives, first for further processing. In this way, it can fill potential gaps in one stream with data received via the other stream.

Market data streams have a highly volatile volume structure. A fast market environment can lead to a self-amplifying effect[6], creating sudden bursts of market data. In liquid instruments, these bursts can happen within a fraction of a second they are called "*microbursts*." The size of a microburst is limited only by the overall processing and delivery capacity of the trading system. This capacity limit can be fairly high with many instruments traded in parallel on distributed systems. But when markets move swiftly, a member doesn't want these high volumes of market data being queued and delayed in their transmission. Trading decisions might otherwise be based on outdated information.

Here are two possible solutions to avoid, or to minimize delays:

1. For very latency-sensitive members—HFT companies and certain algorithmic traders depending on their strategy—a network infrastructure with ample headroom capacity can be used to avoid queuing even during a microburst. The average data transmission rates may be a few Megabits per second. Some members, however, install network connections of 10 Gigabits per second; or even 40 Gigabits per second, i.e., more than a thousand times the average throughput.
2. For many other business models (like screen-based trading), this excessive volume of market data cannot be reasonably processed. Exchange systems therefore offer a "netted" or "pulsed" market data stream. In this stream, order book updates and trades are summarized within a certain time interval. Only the status at the end of the interval is distributed. Sophisticated trading systems permit exchange operators to specify the netting interval separately by traded instrument, or even dynamically depending on the overall volume. The maximum throughput requirements can be better controlled with such netted market data streams. They will not exceed a pre-calculated limit.

In managing bandwidth, IP Multicast has this advantage: Members may individually select certain streams that are essential for their business for subscription. A stream contains the information pertinent to a group of traded instruments. Data transmission via IP Multicast is not 100 % reliable, so exchange systems let members request missing data packages. Alternatively, the system will publish snapshot messages on dedicated streams so that members can reconstruct the order book in the event of gaps in the data received via the normal streams.

Sometimes it is difficult to distinguish between an inactive market and a connectivity issue. This can be the case when members listen to and receive no data in the market data

[6] One order may trigger a cascade of subsequent reactions from other market participants.

stream of a less liquid product. Therefore, exchanges tend to send out "heartbeat messages" on a market data stream. If members receive nothing but the heartbeat, this signals that the market is quiet at that moment. If no heartbeat message is received, a technical alert is raised. It is then obvious that there is a technical issue with the connection, either on the exchange's side or within the member's infrastructure.

7.5 Exchange Connectivity

Exchange organizations need to attract market participants and order flow on a global scale to provide liquidity. Easy, secure, and reliable access for members is fundamental to the exchange business model. A wide range of trading strategies, often requiring different connectivity requirements, may be pursued by members. In response, exchange organizations generally offer a wide range of options for connecting to the exchange system.

The most rudimentary (but sometimes fully sufficient) connection is via the public Internet. Most member firms, however, need a higher level of reliability and guaranteed performance levels. Hence, exchanges often offer connectivity via a dedicated private Wide Area Network, or WAN. Then there are the requirements of technology-driven and latency-sensitive members, HFT traders and many algorithmic traders included. To satisfy this group, exchanges typically also provide co-location facilities as an additional connectivity option. These latency-sensitive trading firms, often connected to multiple exchanges, are willing to pay a significant premium for the fastest connections. Communication technologies, such as *microwave* transmission, are in use in this speed-vital environment.

Connectivity Options: A standard cost-effective way to accomplish direct and simple connectivity is by connecting the member's *front office* to the exchange system via the public Internet. A high level of security can be achieved when using appropriate encryption mechanisms despite the inherently unpredictable nature of Internet data transmission. For small trading firms it is a simpler matter: They may just need a *virtual private network (VPN)* Internet connection and a few standard desktop computers with an Internet browser to easily access multiple exchange systems via the *graphical user interfaces (GUIs).* The GUIs are provided by the exchanges. A VPN is a point-to-point Internet connection through an encrypted data transmission tunnel. It prevents unauthorized third parties from accessing or manipulating data transferred over the Internet.

The reliability and performance of Internet connections cannot be guaranteed because exchange organizations have no control over the Internet infrastructure. These features, however, are critical for the majority of the members.

Business models depend on fast, reliable access to market data provided by the trading system. Hence, exchanges also offer access via dedicated private WANs. These are strictly separated from traffic carried for third parties and protected against unauthorized access. Some exchange organizations offer access

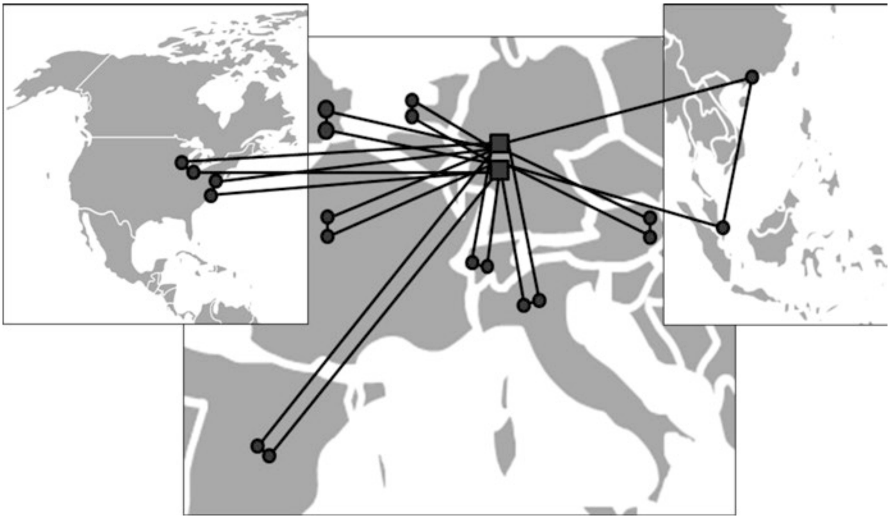


Fig. 7.6 Deutsche Börse's Wide Area Network N7

via specific commercial *extranet* providers; others operate their own global WAN to preserve full end-to-end control between the member installations and the exchange infrastructure.

WANs to connect members with the markets operated by a single-exchange organization are generally built in a star topology. Multiple connectivity centers in different countries and continents—also called points of presence, or access points—are directly linked to the trading system at the center of the star, via the shortest possible path (Fig. 7.6).

Members connect to their closest access point via private telecommunication links, provided by either the member or the exchange organization. Some exchanges ask their members to connect to their connectivity centers, and others provide end-to-end connectivity with options for *redundancy* and bandwidth.

Exchanges develop their trading systems and network infrastructure for full redundancy since reliability is of the utmost importance. The effort and investment in backup infrastructure are substantial. A trading system is typically duplicated, choosing from two options: (1) Both parts are actively used and load balanced over two data centers; for example, the matching engines for one half of the traded instruments are hosted in one data center, and the others are in the other data center. (2) Alternatively, the active primary and the passive backup systems are located in two distinct data centers. Critical data are copied synchronously between the two data centers.

In the event of a large-scale fault in one data center, the installation in the second center will need to take control. For this purpose, exchanges usually select two geographically distinct data center locations to avoid a simultaneous outage in both of them. The cause, for example, could be a regional power interruption, an earthquake, or an extreme weather condition.

Aside from the trading system itself, the network infrastructure and the access points need to be implemented in a fully redundant manner. Access points, similar to data centers, are also installed in pairs of two, and they are interlinked to provide a seamless failover. Each pair of access points is connected via two *backbone* links with the two data centers hosting the trading system. Exchange organizations minimize the risk of a simultaneous outage of both backbone connections. This is accomplished by using different network providers with the highest service level each, whose routes are guaranteed to be physically separate from each other. Sufficient analysis is necessary because seemingly diverse routes can easily turn out to use the same underlying infrastructure, e.g., the same sea cable. A single outage on this infrastructure might then interrupt both supposedly diverse connections. Consequently, a member firm, even a sprawling regional financial community, may be disconnected from the exchange system. It is not so unusual, for example, for the anchor of a fishing boat at sea to cause damage to a major underwater cable[7]. To make things even more problematic, network routes are dynamically altered by the telecommunication providers.

Let me explain: Two routes that have been on separate paths in the past may suddenly share certain underlying infrastructure components after an automatic switch. Hence, the carrier network optimization mechanisms and the routing of individual cables must be verified right down to street level. This will avoid, for example, single points of failure, and unnecessarily long routes. Exchanges monitor network connectivity 24 h per day, enabling them to restore services promptly after a disruption, and to minimize the risk of a complete disconnect or breakdown. In the best-case scenario, this happens before a member would even notice any service degradation.

To offset costs, some exchanges build their own WAN not in a star topology but rather in a ring topology as depicted on the right-hand side of Fig. 7.7. A ring requires fewer backbone connections and less hardware.

In this topology, multiple exchange system locations can be interconnected via a single loop. Every exchange location acts as a connectivity center for all other locations. The members connect to the closest exchange installation. This topology incorporates a natural redundancy because information can flow in both directions around the ring, and a further duplication of links is not required.

However, because an outage of two or more links would impact multiple locations, this topology provides a lower level of redundancy than the star topology. Moreover, network latency in a ring is typically higher than in the star topology. That's because the connection path from a member to the desired exchange installation is on average longer than in the star design.

Algorithmic traders and HFTs create their trading strategies from exchange market data streams, so for them extremely fast processing of market data and equally fast transmission of their order flow are crucial. In fact, low latency is an essential prerequisite for most of these members. Moreover, they must be able to analyze a market situation and react instantly.

[7] Specifically in 2008, a series of sea cable disruptions impacted the data traffic between Europe and the Middle East and Asia.

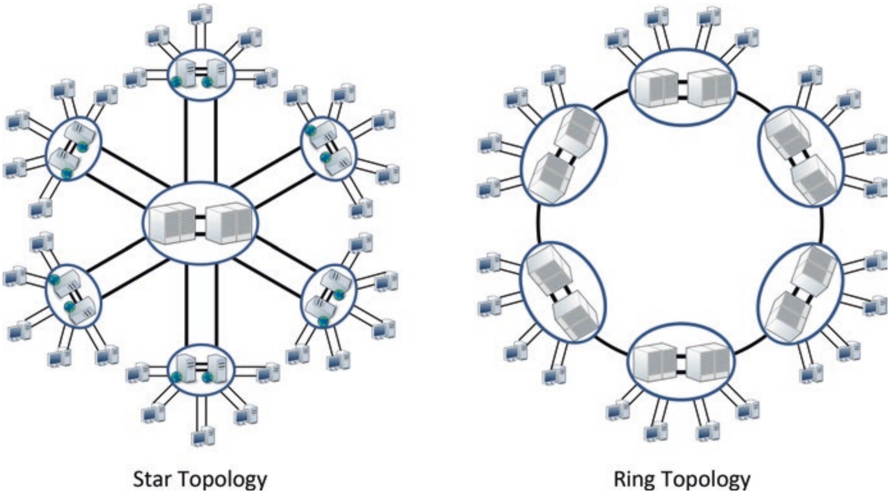


Fig. 7.7 Star network and ring network topology

This is not a new principle. Fast access to market information has always been a key success factor for traders in securities and commodities markets. In the past, timing for traders was a matter of days, hours, minutes, or seconds; nowadays, the time and speed requirements have accelerated, reaching nanosecond turnaround, or a billionth of a second. And so, even the smallest time delay by data transfer from one location to another should be minimized. For the smallest elementary data package, or a bit, it takes 5 μs to travel 1 km through a fiber cable. These 5 μs can make a world of difference; ultimately, this can determine the success or failure of a trading strategy.

Many exchange organizations, seeking to achieve the lowest possible latency for speed-sensitive traders, offer co-location facilities. Under this arrangement, members may install their hardware in exactly the same data center that hosts the exchange back end. Members may then connect locally via so-called cross-connect cables. By minimizing the cable length, it is possible to reduce latency to an absolute minimum, the latency between the member infrastructure and the trading system. Indeed, co-located installations may encounter order round-trip times of approximately a 100 μs—just by cutting out the otherwise inevitable delay from long-distance data transmission.

Some exchanges may also take this step to minimize network hops[8] for these latency-sensitive traders, implementing special high-frequency gateways (see Sect. 7.3) and a dedicated low-latency switching infrastructure. Low-latency switches will use the cut-through technique, which is a method of packet switching. The switch will begin forwarding a packet as soon as the destination address is processed. This method avoids the usual store and forward processing. There is a drawback—relying on the destination devices for error handling.

[8] A network hop represents a networking device on the path between sender and receiver.

This connection concept is highly relevant for certain HFT and algorithmic strategies because co-location installations will always connect faster to the exchange system than any other installation outside co-location. These advantages have lured a large and diverse trading community around specific co-location centers. At Deutsche Börse's co-location center, for instance, more than 150 members are present, including Hudson River Trading, Jump Trading, and Optiver.

As a way to ensure defined service levels between members located in different rooms of the co-location data center, some exchanges use a standardized cable length between the member installation and the exchange infrastructure; others will charge their members contingent on their speed advantage.

In order to limit the impact of a potential data center outage, exchanges play defense, generally preferring to distribute their back-end systems over two redundant data centers. Then there is data transfer and data replication between these data centers. Because it causes additional latency, the trading system infrastructure may be centralized on a single data center campus. Nonetheless, to guarantee the highest possible reliability, a trading system infrastructure would typically be distributed over two separate rooms in the data center. Separate air conditioning and power infrastructure are the ideal arrangement. At the same time, a secondary system must be maintained in a separate, geographically distinct data center to respond to the risk of a complete outage. Data are copied (asynchronously or synchronously) to the secondary data center to allow a market restart after a primary data center failure.

Many latency-sensitive algorithmic traders and HFT firms trade on multiple venues in far-flung global financial centers from New York and Chicago in the USA to London and Frankfurt in Europe—and beyond. Trading strategies on one venue in one city will depend on market data from another venue in another city. With such strategies, speed of data transmission between the market locations has the highest priority. Several competing members will want to be the first to hit an order book.

These market participants are willing to invest in communication infrastructure that allows faster data transmission than the standard telecommunication links between financial centers. They routinely look for ever faster connections between the market back-end locations. A brisk competition for the lowest possible latency has emerged[9]. That has led to some very expensive connection concepts that may deliver speed advantages in the microseconds.

Transmission technologies such as long-distance microwave communication, millimeter waves, and laser links[10] are up to 50 % faster than ordinary fiber cable connections. These speed advantages are directly connected with the physics of light propagation.

---

[9] Some years ago, telecommunication providers started to deliberately construct short cable connections in nearly straight lines of sight between financial centers. That is despite costs being much higher than they would be for standard routing along existing rail lines or highways.

[10] Wireless connectivity options provide faster alternatives in contrast to cable-based connectivity options.

Consider this: While information transmitted via microwaves achieves nearly the speed of light in a vacuum, i.e., 300,000 km per second, data transmission speeds in fiber cable do not exceed 200,000 km per second. Microwave connections are in use between the major market locations in New York and Chicago and between London and Frankfurt. The round-trip time of a microwave connection between London and Frankfurt could, theoretically, be about 2 ms shorter than that of a fiber connection. There are also microwave connections to the landing points of transatlantic cables; however, the idea of installing a series of levitating microwave antennas over the Atlantic still remains science fiction today.

There is a constraint in microwave transmission: It requires straight line-of-sight propagation, and so it relies on a tightly spaced sequence of antennas between sender and receiver. Because the signal weakens rapidly with distance, it needs to be amplified every 50–60 km. Microwave transmissions are also affected by weather conditions and are, therefore, less reliable. The data transfer rates of approximately 150 Megabit per second are also much smaller than in a fiber cable.

Full market data cannot be transferred easily, so members have to diligently filter the most relevant information for transmission. Smaller wavelength, such as millimeter waves, is necessary to increase the bandwidth. Millimeter waves achieve transfer rates of up to 2 Gigabit per second. Unfortunately, millimeter waves must be amplified every 10–15 km because they are even more vulnerable to weather conditions.

The next step to further improve the signal strength and bandwidth would be the data transfer via laser. Test deployment of this is already happening at some highly specialized technology companies.

7.6 Member Infrastructure

Exchange members need to implement a technical infrastructure to connect to the central exchange systems. These infrastructure at member sites vary significantly. They are heavily dependent on members' business models and trading strategies, as well as on their potential customers' requirements.

In the past, many exchanges required that members install special dedicated devices for the particular exchange on their premises (for example to run servers with special software provided by the exchange). The maintenance of these devices would be either the member's responsibility with guidance by the exchange or the exchange would remotely manage the device from their operating centers. A member who connects to several exchanges would have to host and potentially manage a diversified environment of bespoken devices.

State-of-the-art exchange systems nowadays apply the Zero Footprint approach. It is no longer necessary to maintain exchange software at the member site, since the protocols and interfaces to connect to the exchange systems are open and standardized. Instead, members can freely choose suitable hardware and computer operating systems and install their preferred front-end software. In doing so, this may connect to all of the exchange markets that are required by their trading strategy.

Extremely latency-sensitive and technology-savvy members will invest significant effort into creating and optimizing what this software will run on, specifically, the front-end software and hardware platform. These members will typically co-locate their installations at the exchange data centers and, in some cases, they might even deploy specifically designed hardware components such as FPGAs, or self-developed network switches. Others may use third-party software which act as a concentrator for connections to several exchanges. Not surprisingly, there is a most dynamic market for exchange connectivity and order management software.

Front-office software for trading, either custom developed or off the shelf, will receive and display market data with numerous customizable views. Traders can enter, modify, or delete orders for different markets and instrument classes, including basket trades. The front-office software then routes these order messages to the appropriate exchange interface.

Traditional order routing systems forward orders automatically to a predefined exchange. Today's smart order routing mechanisms will flexibly choose to internalize orders, or distribute them between the venues, or forward them to the venue with the best execution capability. Front-office systems increasingly include capabilities for real-time analytics. That allows members to track a trader's performance visually, to set risk limits, and to perform further complex analysis.

Big data, a manifest trend in IT in general, is of particular interest to some short-term investors. The correct investment conclusions from a vast amount of input data can create successful business models in proprietary trading driven by technical market signals. Such algorithmic trading is generally supported by complex, high-performance front-office software. Some vendors provide modular building blocks; in other words, a firm may configure, customize, and run their own algorithms without requiring any special software development skills.

Members active on a variety of market venues will have to diligently design the architecture of the front-office infrastructure. The goal here is huge: The infrastructure should be capable of moving massive amounts of data across the globe, supporting a 24-h trading desk in a follow-the-sun rotation. Regarding performance and latency, it will be critical to select the right geographical location for these front-end installations.

As an alternative, or as a supplement for front-office software, some exchanges also offer their own native front-end GUI. This exchange GUI provides some market data views as well as trading and administrative functions. Workstations that run the native exchange GUI could be connected to the exchange's trading system via the Internet. Yet, they can deliver remarkably good performance by deploying efficient data protocols and transmitting only the stripped raw data.

Alternatively, GUI access can also be implemented over the exchange's private WAN. The exchange GUI solution, as with the other interfaces between the exchange and member, may no longer require that members maintain exchange software at their sites. In some configurations, it is relatively simple: a member only needs a standard desktop computer with an Internet browser and a *Java Runtime Environment (JRE)* to run the GUI. But few members choose an exchange GUI as their preferred solution for actual screen-based trading. Instead, they may prefer to use a software solution with multi-exchange capability depending on the number of exchanges they trade on.

However, the vast majority of members do use the native GUI for other reasons: An exchange GUI may be a sensible choice for a few terminals in a disaster recovery installation, or for an on-site backup. It can also serve as a reference to cross-check the data displayed by the front-office software otherwise used.

Then there is risk management, an increasingly important core component in any front-office system. Agency trading firms as well as proprietary traders will need to control their risk exposure both pre- and post-execution. Therefore, the front-office software will typically connect to a real-time risk management system.

Risk can quickly accumulate and exceed given limits, unless an actual exposure is tightly monitored by an agency trading firm for each downstream client, or by a proprietary trading firm for their own traders. Built-in system safeguards must take immediate action when this occurs. A notable example: The disruption in equity trading caused by a glitch at Knight Capital Group on August 2012 temporarily destabilized trading in nearly 150 New York Stock Exchange-listed stocks. Knight had inadvertently deployed testing software, and consequently suffered a trading loss of US$440 million in less than an hour. This was a reminder that well-designed risk management safeguards are essential.

Robust risk management also requires certain *post-trade* functionality. Once a trade has been executed, it will be reported by the clearing house, either to the member's *back office* or to the designated clearing firm. In the latter case, it is presumed that the member has a clearing arrangement with a partner. *Middle-office* and back-office support is typically provided by one of a few market-leading software solutions.

This post-trade functionality is basically straightforward, yet very critical. Post-trade facilities maintain and manage aggregated positions, and provide the tools to assess underlying exposures for an individual book or across multiple instruments. Moreover, post-trade facilities analyze positions per trader, or for a trading desk, even for an entire firm.

7.7 Time Management and Performance Monitoring

Exchange organizations must provide full transparency for each single transaction due to their special economic significance and major financial impact. In fact, many members expect an exchange system to provide information about the exact point in time a message hits the exchange. More precisely, they expect to know when it enters the exchange gateway, and which subsequent chain of events will be triggered and when.

The best way to provide this kind of transparency is to *time stamp* every message at each step of its path through the electronic trading system. Naturally, this only makes sense if the clocks of the different devices in the system have exactly the same time and run at the same speed. Not surprisingly, the clocks need to be synchronized very frequently to ensure that this is so.

In general, this is accomplished by using a network protocol, such as the traditional *Network Time Protocol (NTP)* for clock synchronization between devices in a computer network. Time, as provided by a reference clock, is being propagated throughout the network. An accuracy rate of approximately 1 ms that can be accomplished with NTP is not necessarily sufficient for low-latency trading systems.

Exchange organizations must be able to handle fast-moving markets. Time resolution in the sub-microsecond regime is required. Moreover, the electronic exchange system itself is a highly complex system, so a synchronized time signal throughout this system all the way down to member installations is desirable. These requirements can only be met with a more sophisticated time management protocol.

To that end, the *Precision Time Protocol (PTP)* which is typically used is able to handle hundreds of servers, achieving a much higher level of accuracy than the standard NTP. Exchanges deploy specific hardware timing components to achieve extremely high accuracy within the exchange infrastructure, and the member co-location installations. A single, highly precise reference clock is the sole source for time synchronization. This clock will typically use the *global positioning system (GPS)* signal; it provides accuracy to a fraction of a microsecond.

Still, because exchange infrastructure is highly critical, one may not want to depend exclusively on the GPS. A standard radio time signal could be used as well. The radio time signal would serve as a reference and backup, in case the GPS signal is lost or may have been manipulated.

Time protocols measure the delay caused by information transfer between devices. They are, therefore, able to propagate the appropriate time within the network. The transfer time is calculated by averaging the forward and the return time (Fig. 7.8).

The calculation of the signal delay

$$\Delta t = \frac{(t_2 - t_1) + (t_4 - t_3)}{2}$$

assumes that the forward and the return times are equal. In practice, this is often not the case. One major problem: potential queuing in the timing devices during high workload. Such queues cause delays and differences in the transfer times, a

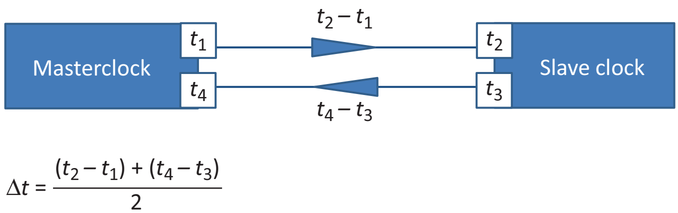


Fig. 7.8 Averaging delay when synchronizing time in a geographically dispersed installation

problem that can be solved by hardware time stamping. The hardware does the time stamping as soon and as fast as possible after the arrival of a message, and as late and fast as possible before the message leaves, a process which avoids software queues.

Devices in the exchange environment regularly receive a precise time. Still, their own systems' internal clocks may still have a slight drift, i.e., a bit too slow or bit too fast. When the device receives the next precise time, it will have to adjust its own system clock accordingly. This adjustment can be done in two different ways:

- By abruptly jumping to the right time, which adjusts the clock instantaneously, but may cause a shift in the chronologic order of the specific device.
- In the form of a smooth and gradual convergence. That means that it takes more time to adjust the clock but the chronologic order of the specific device is conserved.

The second approach to synchronization is preferred because, for exchange organizations, chronological order is highly important.

Exchanges and their members can assign precise time stamps to messages at crucial processing steps based on very accurate time synchronization. Hence, time stamps on these servers can be used to analyze one-way transport times. Figure 7.9 diagrams a typical example of a member sending an order request message, and being answered by a private order response message and a public order book update.

Figure 7.9 can be interpreted as follows:

- The time stamp t_1 can be taken by the member application when the request is sent.
- t_3 is taken by the exchange gateway when the request is read on the member's side of the gateway.
- t_5 is taken by the exchange matching engine when the request is read there.
- t_7 is measured at the time when the matching engine maintains the order books.
- t_6 is taken by the exchange matching engine when the response is sent from the matching engine to the gateway.
- t_4 is taken by the exchange gateway when the response is sent from the gateway to the member.
- t_2 can then be taken by the member application when the response is received.
- t_8 is taken by the market data interface, before the information is sent to the member.
- t_9 can again be taken by the member application when the respective market data arrives.

Only time differences like (t_2−t_1) can be analyzed in case of non-synchronized times. That is because discrepancies in absolute clock times are eliminated by taking the difference.

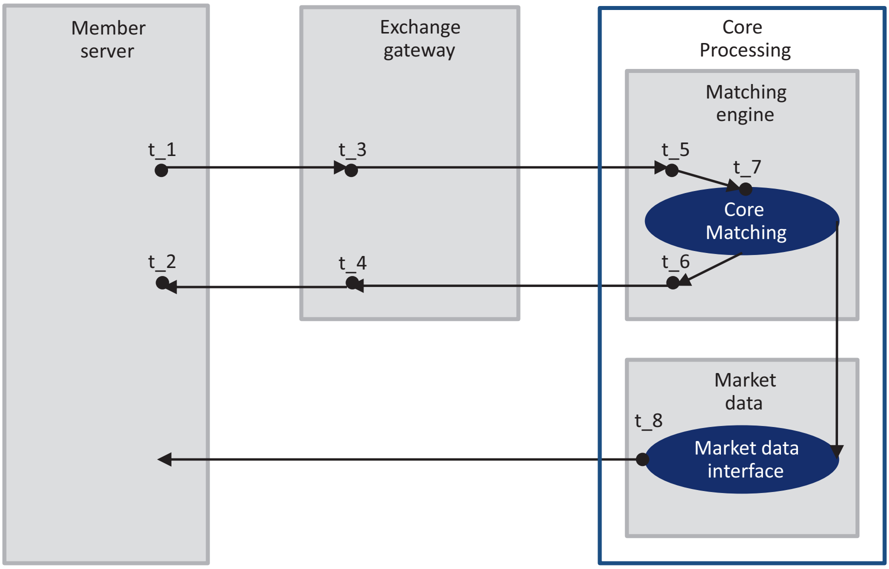


Fig. 7.9 Time stamps in the order processing event sequence

Exchanges may use the above as well as additional time stamps to offer advanced trade traceability to their members. When match occurs in the exchange order book, member order request messages trigger further conditional messages. Examples are order event and trade confirmation messages. By time stamping these downstream messages, and linking them to their parent messages via unique identifiers, exchanges can build entire message trees. The exchange can track the complete life cycle of a message and subsequent events in this way. Intelligent assignment mechanisms make it possible to add these time stamps with minimal impact on overall performance and latency.

The technical support staff at an exchange, with this complete data history, can conduct detailed performance analysis, troubleshooting, as well as capacity management. And because some or all of these time stamps are also available to members, there is full end-to-end transparency. That means that trading firms may analyze system behavior and optimize their infrastructure accordingly.

7.8 Conclusion

In this chapter, we have seen that today's equity markets depend on state-of-the-art information technology. A fully electronic trading environment must balance competing objectives, including reliability, transparency, and high performance.

The exchange needs to account for and to reconcile a diverse set of technical and functional requirements within the exchange member community. The expectations of latency-sensitive market participants have proven to be the strongest driver for innovative and pioneering technology concepts in equity trading systems.

The next generation of technology will continue to transform the exchange system architectures and the exchange ecosystems as a whole. Blockchain technology, cloud computing design principles, big data processing, and mobile computing, to name a few, will create new unprecedented opportunities to shape the future of the financial industry.

Further Reading

St. Hammer, *Architects of Electronic Trading: Technology Leaders Who Are Shaping Today's Financial Markets*, Wiley, 2013

Trading Technologies Ltd., *A Strategic Study of Stock Exchange Technology*, Metal Bulletin, 2003

C. Zaloom, *Out of the Pits: Traders and Technology from Chicago to London*, University of Chicago Press, 2006

D. Lohfert, "Battling the bottleneck: an analysis of the new Eurex platform", *Automated Trader Magazine*, Issue 30, Q3 2013, pp. 16–20. Also available at: http://www.automatedtrader.net/articles/analysis/144302/battling-the-bottleneck-an-analysis-of-the-new-eurex-platform.

H. Cumberland, 2014. Futures & Options World: "Riding along on the crest of a wave". Available at: http://www.fow.com/3308917/riding-alon.html.

St. Hammer, 2014. Futures & Options World: "The 'Race to Zero' continues with faster hardware". Available at: http://www.fow.com/3315483/The-Race-to-Zero-continue.html.

D. Wigan, 2013. Futures & Options World: "Microwave arms race gathers speed". Available at: http://www.fow.com/3248074/Microwave-arms-race-gathers-pace.html.

Chapter 8
Contractual Relationships Across the Value Chain

Matthias Stötzel

Three layers must be carefully considered in describing an equity market's value chain, from processing orders for trading at an exchange to the clearing and settlement of transactions: the trading layer, the clearing layer and the settlement layer. These layers are technically integrated by straight-through processing (STP)[1] of transactions. In this chapter, however, we will focus on the legal level, i.e. the regulatory and contractual relationships that are required. At minimum, these are the relationships that are usually established on each layer to execute and process securities transactions along the value chain.

8.1 The Trading Layer

As already outlined in Chap. 2 legal relationships on the trading layer are established by the exchange with issuers of securities that are permitted to trade on the exchange and with trading participants. Before we take a closer look at these legal relationships (Sects. 8.1.2 and 8.1.3), we briefly describe in Sect. 8.1.1 the role of the securities exchange and of issuers and trading participants.

[1] Straight-through processing (STP) enables the entire trade process for capital market and payment transactions to be conducted electronically without the need for re-keying or manual intervention.

M. Stötzel (✉)
Deutsche Börse AG, Mergenthalerallee 61, Eschborn 65760, Germany
e-mail: Matthias.stoetzel@deutsche-boerse.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_8

8.1.1 Exchanges, Issuers and Trading Participants

The securities exchange is an organised market, bringing together supply and demand for securities. In this role, it is part of the secondary market for trading securities after the securities are issued on the primary market, a task typically accomplished through investment banks. The issuance of new shares is referred to as initial public offering (IPO), a company having previously been in private hands going public so that the new shares become tradable. This assumes that the requirements in the statutory law and in the exchange rulebook are fulfilled. The benefits can be enormous. An IPO can increase awareness of a company and its products on a national and international scale, depending on the exchange selected by the company as its listing venue. Furthermore, an IPO strengthens the equity base and enhances the creditworthiness of a company, therefore contributing to the company's future prospects and competitiveness. The primary market is described in more detail in Chap. 3 of this book.

On the one hand, securities exchanges, as part of the secondary market, offer investors the opportunity to participate in the economic progress of corporations that have issued securities listed on the exchange. We refer to trading participants as investors since trading participants (dealers or principals) may bring orders to the exchange for their proprietary business, or on behalf of customers (as brokers or agents). On the other hand, securities exchanges fulfil a central function for companies demanding capital for funding their business operations. For this purpose, they issue securities to be traded on the exchange.

The same is true for the capital requirements of the state and state organisations. Bringing together investors with issuers of securities, exchanges fulfil a central function in the economy. As mediator, exchanges make possible an optimal and efficient reconcilement of interests between these groups. These same groups would otherwise have to negotiate directly at considerably higher cost in the form of search and information costs. The role as mediator indicates that securities exchanges (different from central counterparties, CCPs, which are contractually interposed between the trading participants) are neither party to transactions executed on the exchange nor hold or own securities.

The securities exchange's main responsibility is fair and orderly price discovery for issued securities accepted for trading. Prices determined by the execution of orders have to best reflect the investors' desire to buy and to sell securities and the current value of the securities. Price determination and the execution of orders can take place in different trading models, from continuous trading and periodic call auctions to market maker-driven trading and other models. Securities exchanges, generally, facilitate trading in more than one trading model, depending on the character and liquidity of securities. We turn to the process of price determination in more detail in Chap. 4 of this book.

Without labouring the details, two aspects are worth considering as background for the legal relationships established by the securities exchange with issuers and trading participants.

- First is the exchange ownership structure. Historically, most exchanges were non-profit organisations owned by their members. However, many exchanges have transformed from a traditional mutual or cooperative organisation into a for-profit, shareholder-owned company, a change referred to as demutualisation. Some exchanges, like Deutsche Börse and the New York Stock Exchange (as part of the Intercontinental Exchange today), have become public companies listed on their own markets because of demutualisation. Other exchanges have demutualised but are still private corporations, their previous members remaining owners. Moreover, demutualisation may influence the strategic playground, opening up more freedom for exchanges to strategically develop exchange trading and to introduce new types of services. Still, in a demutualised ownership structure, divergent interests of securities exchanges and their trading participants and issuers tend to be more pronounced.
- The second aspect is the power of securities exchanges to set rules for exchange trading and the settlement of exchange transactions. Usually, this is referred to as self-regulation. The term is misleading insofar as exchanges only regulate themselves to a limited extent, particularly in matters of exchange organisation and governance.[2] The main focus of self-regulation, however, is the regulation of the exchange-related activities of issuers and trading participants by the exchange. This means that the legal relationships to issuers and trading participants are to a large extent subject to self-regulation by the exchange. The power of self-regulation is delegated to exchanges by the legislator. Therefore exchanges, exercising this power, are obliged to observe the statutory provisions of this legal privilege as well as the instructions of the regulators responsible for the supervision of exchanges.

Regulation on the trading layer, based on this concept, is a balance between self-regulation by the exchange close to the their trading participants and issuers and governmental regulation that ensures orderly exchange trading and settlement of transactions, including equal treatment of issuers and trading participants. From the issuers' and trading participants' perspective, both are subject to exchange rules. Depending on how the concept of self-regulation is put into action, issuers and trading participants may be represented, as is the case in Germany,[3] in the rule-making bodies of the exchange. The effect is that trading participants and issuers are participating in the rule-making process of the exchange. In this approach, they are able to decide on the strategic development of the exchange, irrespective of whether they are owners of the exchange, or the exchange has demutualised (Fig. 8.1).

---

[2] In Germany, legal framework for the operation of exchanges is provided for in the Exchange Act. Under the Exchange Act, power is delegated to the exchanges to issue rules, e.g. on organisation of the exchange, trading models and settlement of exchange transactions. Such rules require approval of the stock exchange supervisory authority.

[3] In Germany, the exchange council is responsible for the issuance of the exchange rules. Pursuant to the German Exchange Act, trading participants and issuers must be represented in the exchange council.

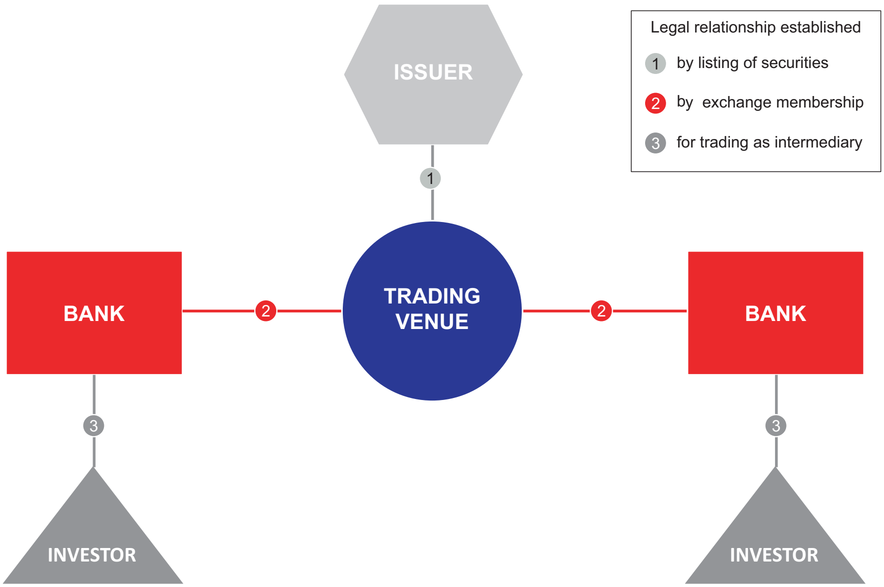


Fig. 8.1 Legal relationships on the trading layer

8.1.2 Legal Relationship to Issuers

Products traded on the securities exchange, as part of the secondary market, are securities issued on the primary market by companies that we refer to as issuers. We will concentrate on private companies as issuers with which legal relationships are established by the exchange though also government bonds may be traded on the exchange.

8.1.2.1 Balance Between Issuers' Interests and Investor Protection

The issuers applying for permission to list and trade on the exchange primarily are motivated by raising capital separately from financing by banks. Issuers can choose from two sources of capital: equity capital by issuing shares, i.e. with an IPO via the exchange, or debt capital by issuing bonds. Issuers are subsequently able to obtain additional capital by means of capital increase,[4] or by issuing more bonds to finance growth, or to optimise their capital structure, and sometimes both. Consequently, issuers have a higher degree of independence and are enabled to raise capital cost-effectively.

[4] Capital increase is a method used by corporations to raise share capital by giving existing shareholders the right to subscribe to new shares for cash. Alternatively, capital can be raised by exchanging assets such as shares in another company or by raising the par value of existing shares.

Exchanges, nonetheless, have to create an environment attractive enough for issuers to choose as a listing venue. At the same time, exchanges must be careful that the capital providers—investors—are sufficiently protected, and in particular that they have reliable information about the issuers' business to make informed investment decisions. Against this background, the legal relationships with issuers established by the exchanges within their rules and the framework of statutory provisions essentially are about striking a proper balance between the interests of issuers and investors. In particular, this means setting and enforcing rules for the admission of securities to trading on the exchange and for the post-admission obligations of issuers.

8.1.2.2 Admission of Securities to Trading

The admission of securities for trading publicly on an exchange is granted by the exchange subject to the application of the issuer. The issuer has to submit information and documents to provide evidence that the admission requirements as mentioned below are fulfilled. The application may have to be filed by the issuer together with a bank or an investment firm. This is to ensure that the issuer is provided professional support and the issuer is aware of, and able to meet, the admission requirements as well as post-admission obligations. Once admitted, a legal relationship is established between the securities exchange and the issuer, and is based on the applicable statutory law and exchange rules. The exchange rules here must be clear and transparent, and ensure that securities admitted are capable of being traded in a fair and orderly manner. The following is a required checklist:

- The incorporation of the issuer complies with applicable law.
- The securities are standardised and freely tradable (fungible).
- The securities possess an International Securities Identification Number (ISIN).
- The securities have an expected market value of a specific minimum amount.
- The orderly fulfilment of the transactions in the securities is guaranteed. (That requires that the securities are eligible for clearing as well as custody and settlement as described in Sects. 8.2 and 8.3.)
- There are no regulatory prohibitions against the trading of the securities on the exchange.

Once the securities are admitted, the issuer has to continuously fulfil certain requirements, besides the post-admission obligations as described in (c) below. The exchange, aiming at the protection of the trading participants, has the responsibility to review regularly the compliance of the issuer with the admission requirements for the securities admitted. Basically, the enforcement of requirements—set out in the statutory law and not the rules of the exchange—is not the responsibility of the exchange but of the state regulator. Once the admission of the securities and the legal relationship with the issuer are established, the exchange has to ensure that the issuer has the right for the securities to be traded on the exchange. The exchange may suspend trading of the securities if orderly trading on the exchange is temporarily endangered, or if the suspension is required to protect investors.[5] What happens if orderly trading on the stock exchange no longer appears to be ensured? As an instrument of last resort, the exchange may discontinue trading of the securities and revoke the admission of securities to trading on the exchange. This may also apply if the issuer does not meet its obligations under the terms of the admission. Furthermore, the issuer can apply for (and the exchange can approve) the revocation of the securities admitted, unless the application conflicts with investor protections. To that end, exchanges may foresee that, for example, if the securities are not traded on another market, trading of the securities on the exchange will be discontinued only after an appropriate period of time has passed. This will allow trading participants to sell the securities. The legal relationship between the exchange and the issuer is then terminated with discontinuation of trading and revocation of admission.

In exceptional circumstances, if securities can be admitted to trading by the exchange without the issuer having applied for admission (so-called ex officio), no legal relationship between the exchange and the issuer is established. Issuers with securities admitted to trading without application, i.e. without the issuer's consent, may not be obliged through the rules of the exchange to fulfil post-admission obligations.[6] Usually, the exchange will inform the issuer about the admission of the securities.

8.1.2.3 Post-admission Obligations of Issuers

The issuer will be in the public limelight more than ever before, once shares of this newly public company are admitted for trading. And while this leads to the opportunity of raising capital separately from financing by banks, the issuers also have a challenge, fulfilling certain post-admission obligations, aimed at supplying information to investors about their business. Investors can then analyze and manage their investments on the basis of sound and reliable information. These obligations only apply to securities admitted to trading, and are embedded in the legal relationship between the issuer and the exchange. They are set forth, nevertheless, in statutory law as well as in the rules of the exchange. Basic requirements must be fulfilled under statutory law irrespective of the exchange on which their securities are traded. In addition, exchanges in their rules may foresee further post-admission obligations of issuers and, consequently, establish listing segments with different levels of transparency for investors.

---

[5] For example, trading of securities will be suspended if orderly determination of exchange prices or the orderly settlement of exchange transactions is not ensured.

[6] Pursuant to Section 48 (1) German Exchange Act, issuers whose securities were listed on the exchange without their consent may not be obligated through trading guidelines to publish information in regard of these securities.

Statutory law provides a set of obligations for the issuers deemed essential for the dissemination of information to the public, and for the avoidance of market manipulation. One important component in these obligations is so-called inside information. That can be defined as any specific information about circumstances which are not public knowledge relating to an issuer of listed securities, or to the securities themselves. The idea here is that if this information became publicly known, it would likely have a significant effect on the market price of the listed security.

Issuers are under the obligation without undue delay to publish all inside information that directly concerns that issuer (so-called ad hoc publication). Prior to publishing inside information, in certain circumstances,[7] the issuers are obliged to notify the exchanges where the securities are traded. The exchanges can assess if orderly trading on the exchange is endangered based on this information, and if securities trading has to be suspended or even discontinued.

Furthermore, issuers of listed securities have to maintain insider lists, i.e. lists of persons working for them who have access to inside information as part of their function. In general, it is widely prohibited—and not only for issuers and their management and staff—to make use of inside information in these circumstances: to acquire or dispose of listed securities for oneself or for a third party, or to disclose or make available inside information to a third party without the authority.

Furthermore, issuers of listed securities by statutory law are obliged to publish information in financial reports on a regular basis. These reports in particular have to include financial statements as well as a management report.

The extent to which securities exchanges in their rules may provide for additional obligations is driven by the appropriate statutory law. If a comparatively high standard is provided for on the level of statutory law, there is only little room for manoeuvre on the exchange level. In contrast, if statutory law requires only a lower standard, exchanges may create listing segments by imposing post-admission obligations for issuers.

Exchanges may set up listing segments with different levels of post-admission obligations based on the size and business of the issuers, and within the scope of self-regulation. In general, exchanges will impose further and more stringent obligations for large- or medium-sized and internationally operating issuers, than for young, growing, small- and mid-sized companies. For example, exchanges may require issuers to regularly conduct analysts' meetings, and to prepare and update a financial calendar with details on the most important corporate action events of the issuer.

Exchanges may also introduce listing segments tailored to these securities to promote trading of certain securities. In these segments, issuers are required to observe especially high transparency requirements. These requirements must be met both at the time when the issuer is first admitted to the listing segment and on a

---

[7] Pursuant to Section 15 German Securities Trading Act, before publication of inside information, the issuer shall notify the management of the exchange on which the securities are admitted to trading.

continuous basis afterwards. For instance, an exchange may require issuers of bonds traded in a certain listing segment to submit and regularly update a bond rating and company profile to increase transparency for investors, with a view to the valuation of the bonds issued.

8.1.3 Legal Relationships to Trading Participants

Legal relationships between the securities exchange and trading participants are the result when companies are admitted to trade on an exchange. As is the case for issuers, the legal relationships are based on statutory law as well as on the rules of the exchange.

8.1.3.1 Trading Participants, Intermediaries and Investors

The term "trading participants" refers only to those companies that are admitted to trading on the exchange. They do not represent exclusively but are part of the group of institutional and retail investors who are from an economic perspective (as principals) the basic source of order flow on the exchange. When they are acting on their own account as principal, trading participants can be considered investors. In contrast, when they are trading securities on behalf of customers, trading participants are acting as intermediaries for investors.

The investors may be other companies including mutual funds, pension funds and insurance companies. They may also be individuals (retail customers) who either do not meet the requirements for the admission to trading on the exchange or for business reasons have decided not to become trading participants. (Individuals, in general, do not meet the requirements for the admission to trading on the exchange.) As a rule, exchanges are only establishing legal relationships to trading participants but not to investors who are not admitted to exchange trading. The rights and obligations of these investors are the result of contractual relationships they have established with trading participants (acting as intermediaries). To a certain extent they reflect the rights and obligations of trading participants with the exchange.

8.1.3.2 Admission of Trading Participants

In exchange trading, the participation of companies and individuals (exchange traders) who are acting on behalf of the trading participants requires admission from the exchange. Once approved for trading on the exchange, legal relationships between the exchange and trading participants and exchange traders are established based on statutory law and the exchange rules. The filing of applications already establishes, procedurally, legal relationships, based on which applicants have to furnish evidence, demonstrating that conditions for admission are met. The exchange is under the obligation to process the application in compliance with the applicable law and exchange rules.

To obtain admission to trading on the exchange, both the requirements of statutory law and of the rules of the exchange must be fulfilled. These requirements shall ensure that companies and individuals have the qualification to participate in orderly exchange trading, and guarantee the orderly settlement of exchange transactions. In case of transactions cleared through a CCP, the CCP will become the contracting party of the trading participants. For transactions that are to be cleared bilaterally between the trading participants, the fulfilment of the admission requirements further ensures that trading participants will only conclude transactions with suitable counterparties. Usually, companies admitted to trading on the exchange must fulfill the following:

- Conduct purchasing and selling of securities for their own account, or for the account of third parties or brokerage business with a commercially organised business establishment.
- The persons entrusted with managing the company's business and authorised to represent the company must be reliable and have the professional qualification necessary for participation in exchange trading.
- The orderly settlement of transactions on the exchange is ensured, which requires that the company conducts the settlement of transactions through a Central Securities Depository (CSD) recognised by the exchange, and, in case of transactions cleared through a CCP, ensures clearing via this CCP (we will discuss details in Sect. 8.2).
- The company provides evidence of equity capital as determined in the exchange rules, or by statutory law; and there are no facts justifying the assumption that the company, taking into account the equity capital evidenced, does not have the necessary economic capacity to participate in an orderly manner in exchange trading.
- The company fulfils the technical and legal requirements to access the trading system of the exchange.

Access to the trading systems of the exchange and fulfilment of the requirements by the trading participants are of paramount importance. This enables trading participants to trade on the exchange, that is, to enter orders or quotes into the trading system, and to modify or delete them as well as to receive data from the system. Usually, access to the trading system is based on agreements between the exchange and the trading participants. To that end, the General Terms and Conditions[8] lay out the rights and obligations of the parties and the technical details for the connection of the trading participants' technical infrastructure to the trading system of the exchange. These agreements may include

[8] General Terms and Conditions are general and special arrangements, provisions, requirements, rules, specifications and standards that form an integral part of an agreement.

explicit provisions concerning the conditions trading participants may be entitled to claim in case of disruption or default of the trading system.[9]

Exchanges may include co-location services as part of these agreements, allowing trading participants, especially algorithmic traders,[10] to place their trading engines in the exchange's data centre. By shortening the distances between the trading participants' trading engine and the trading system of the exchange, this service allows latency-sensitive trading participants, such as high-frequency traders, to shorten round-trip times in trade executions—and with lightning fast access to exchange trading data. Based on the principle of equal treatment, exchanges are obliged to offer co-location services to all trading participants on equal terms. The decision to avail of these services is the sole responsibility of the trading participants, and is based on their business models.

The companies applying for admission to trade must identify the individuals (exchange traders) who plan to participate in exchange trading on behalf of the companies. These individuals also require admission by the exchange on the condition that they are reliable and possess the professional qualification. Professional qualification is awarded if the persons have the expertise and practical knowledge for trading on the exchange. This expertise may be demonstrated by a successful participation in an exchange trader examination if it is offered by the exchange.

The exchange has to continuously monitor trading participants and exchange traders for compliance with the trading admission requirements. The trading participants and exchange traders, in turn, must inform the exchange about any changes that would lead to a loss of their admission. The exchange may suspend or revoke—if non-fulfilment is not temporary—the admission of trading participants or exchange traders to trading on the exchange based on such information. As a result, the legal relationship between the exchange and the trading participants or exchange traders will be suspended or terminated.

8.1.3.3 Agreements with Market Makers

Securities exchanges commission trading participants as market makers to increase liquidity, and to ensure the quality of price determination. Depending on the supported trading models, and the type and liquidity of securities admitted to

---

[9] These connection agreements may further provide for details of the technical connection (access alternatives, interfaces and specifications), requirements to be met by the trading infrastructure of the trading participant, unilateral amendments to the agreements made by the securities exchange, confidentiality obligations, governing law and place of jurisdiction, termination as well as fees for the technical connection (by way of a price list incorporated into the agreement).

[10] Algorithmic trading means trading in securities where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price or quantity of the order or how to manage the order after its submission, with limited or no human intervention.

trading, commissioning market makers usually is based on separate agreements between the exchange and the market makers. These agreements, in particular, spell out the securities that market making services will be provided for as well as the compensation of the market makers in the form of commissions, or rebates on trading fees. (Further revenue sources of market makers are profits from the bid-ask spread of quotes provided by the market makers.)[11] Each security may be assigned by the exchange to several market makers or exclusively to one market maker. Exclusivity applies particularly when certain market maker functions are required for specific securities, or because of the securities type. In some instances, exchanges cannot expect to be competitively successful in attracting a market maker without granting this exclusivity. In this case, if more than one trading participant is interested in performing the market maker function, the exchange must have transparent and non-discriminatory rules in commissioning a market maker for each security.

In addition to the agreements with market makers, the exchange rules often outline the requirements to perform market maker functions. These include availability of sufficient personnel, technical and financial resources as well as relevant expertise and experience. Furthermore, the basic tasks and obligations of market makers are set forth in the rules of the exchange. In particular, these include the obligation of market makers to enter binding quotes into the trading system, i.e. the parallel entry of a limit buy order and limit sell order. The exchange rules may also establish certain requirements for market makers providing quotes. For example, these include maximum bid-ask spreads (i.e. difference between bid and ask side of a quote); minimum quote volumes; and certain minimum times that quotes must be available during trading on the exchange. These requirements are designed to ensure that market makers perform their function—providing additional liquidity, particularly for less liquid securities. Securities with lower liquidity can also be traded continuously on the exchange with the support of market makers. It is the task of the exchange to monitor and enforce the fulfilment of market maker requirements. If a market maker repeatedly does not meet these requirements, the exchange may terminate the market maker agreement.

We have referred above only to market makers who are under the obligation to continuously provide quotes. Exchanges have also established market maker models based only on financial incentives for trading participants granted by the exchange if certain requirements are fulfilled. As in these market maker models trading participants are not contractually committed to enter quotes into the trading system, exchanges cannot rely on the provision of additional liquidity.

[11] Market maker agreements usually are setting forth the tasks of the market maker and requirements to be met by the market maker (reference will be made to the exchange rules insofar as such rules contain corresponding provisions), the requirements for the inclusion of securities into the agreement, unilateral amendments to the agreements made by the securities exchange, confidentiality obligations, governing law and place of jurisdiction, termination as well as fees or rebates to be paid by the exchange to the market maker.

8.2 The Clearing Layer

Transactions executed on the securities exchange enter the clearing process and, therefore, details of the transactions are passed to the CCP. All larger securities exchanges have appointed one or more CCPs to clear transactions. Exchanges may foresee a bilateral settlement of transactions in certain securities. This is the case, most notably, if securities, due to insufficient liquidity, are ineligible for clearing through a CCP. In this instance, exchanges in their rules provide for the settlement obligations (delivery and payment) of the trading participants.[12]

In this section we will concentrate on clearing transactions through a CCP. As background, we describe the relevant players on the clearing layer (Sect. 8.2.1), the CCP role and the transactions concluded in central clearing (Sect. 8.2.2) as well as the admission of clearing members (Sect. 8.2.3). Thereafter, we will discuss the legal relationships established by the CCP with clearing members and non-clearing members on the basis of bilateral and tripartite clearing agreements and further agreements accessory to these legal relationships (Sect. 8.2.4).

8.2.1 CCPs, Clearing Members and Non-clearing Members

The first step to understand clearing of transactions through a CCP is to become familiar with the players. These players are also referred to in the exchange rules that set the requirements for the orderly settlement of transactions. Unless transactions are settled bilaterally, trading participants must ensure clearing via the CCP, pursuant to these provisions. Trading participants can meet this requirement in two different ways, either by a direct membership in the CCP (as a clearing member) or an indirect connection (as a non-clearing member) through another party who is a clearing member. In the former, companies are not only exchange members but are also members of the CCP. In the latter, formal agreements between non-clearing members and their clearing members are a condition for admission as an exchange trading participant.

According to the rules of the CCP, or clearing conditions, clearing membership may come in two categories depending on the transactions clearing members are permitted to clear. A general clearing member (GCM) may clear the transactions of customers, non-clearing members and its own transaction. A direct clearing member (DCM) is permitted to clear only transactions of its customers and company-affiliated non-clearing members, and its own transactions. Thus, direct clearing members are allowed to clear only for non-clearing members that are part of the same company group, but not for other non-clearing members. Clearing members

[12] For bilateral settlement of securities transactions, the rules of the exchange may in particular provide for the settlement obligations of the seller and the buyer, the time of settlement and the procedure to be observed in case of late settlement.

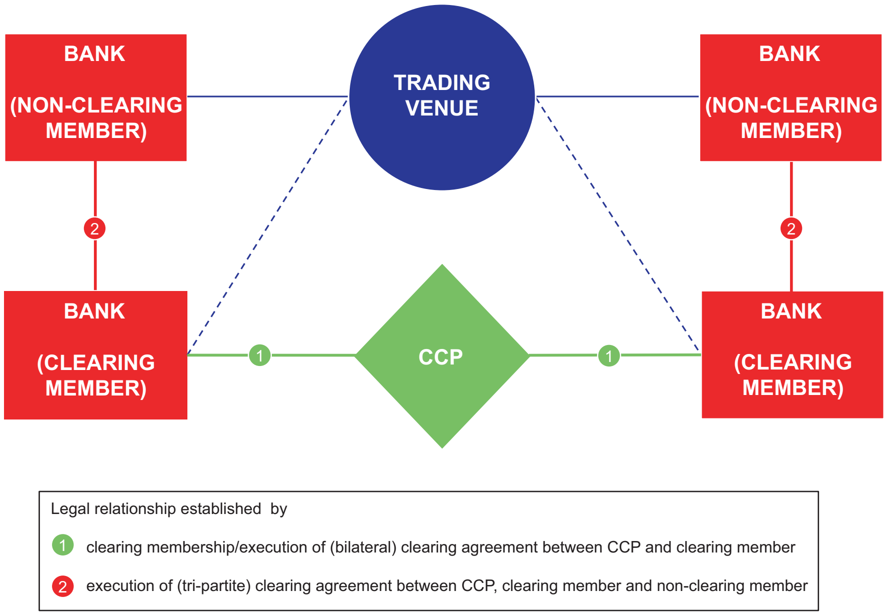


Fig. 8.2 Legal relationships on the clearing layer

guarantee the delivery of securities, and are responsible for the cash settlement of transactions. They are required to provide the CCP with their own margin positions as well as the positions of their clients and non-clearing members. Non-clearing members have limited access to clearing functions, and make margin deposits to the CCP via the contractually stipulated clearing member (Fig. 8.2).

8.2.2 The Role of the CCP and Transactions in Central Clearing

The CCP clears transactions in securities eligible for central clearing that result from matched orders or quotes of trading participants of the securities exchange. In so doing, the CCP becomes a seller for each buyer and a buyer for each seller. In the role of the seller of securities, the CCP is obliged to deliver the respective securities to the buyer; in the role of the buyer it is obliged to pay the agreed price for the delivered securities. The details of the delivery and payment obligations are outlined in the clearing conditions of the CCP.[13]

[13] Comparable to exchange rules providing for a bilateral settlement of securities transactions, the rules of the CCP are setting forth the settlement obligations of the clearing members involved as seller and buyer to a transaction, the time of settlement and the procedure to be observed in case of late settlement (buy-in and cash settlement).

A direct transaction-based contractual relationship is established by the CCP only with clearing members. Therefore, the CCP obligations resulting from transactions executed on the exchange exist only vis-à-vis clearing members. If non-clearing members are involved, similar transactions are established between the clearing and the non-clearing members with the interposition of the CCP. Accordingly, in this case securities transactions exist between the following:

- Non-clearing member A (as seller) and clearing member B (as buyer)
- Clearing member B (as seller) and the CCP (as buyer)
- The CCP (as seller) and clearing member C (as buyer)
- Clearing member C (as seller) and non-clearing member D (as buyer)

The interposition of the CCP between the buyer and the seller of securities can be based on two different legal concepts. In the "open offer" concept, the CCP rules and the corresponding rules of the securities exchange permit the CCP to make an open offer to all clearing members, and to enter into a transaction as soon as orders or quotes of trading participants are matched in the trading system of the exchange. Consequently, no transaction is concluded directly between the trading participants at any point. In contrast, in the "novation"[14] concept through matching of orders or quotes on the exchange, a transaction between the trading participants is concluded. Only then, after certain checks by the CCP are concluded, the CCP by use of novation interposes itself between the clearing member as buyer to the transaction and the clearing member as seller to the transaction. In this manner, the initial transaction between the trading participants is replaced by similar transactions between the CCP and the clearing members (and between clearing members and non-clearing members). In practice, there is no difference between open offer and novation as long as novation actually takes place for all transactions. If this is not the case, and therefore no transactions with the CCP—and between clearing members and non-clearing members—are established, it must be determined by the trading participants whether the transactions between them shall be upheld or terminated. That scenario cannot occur in the offer concept that is resulting in transactions with the CCP once orders or quotes of trading participants are matched on the exchange.

In the event of the failure of any clearing member to fulfil its obligations resulting from transactions, the CCP assumes responsibility, vis-à-vis the other clearing member involved in a securities transaction. To mitigate the risk of a clearing member failing, the CCP holds margin in the form of cash and securities against open positions from its clearing members. Furthermore, only companies that fulfil certain admission requirements are entitled to participate as clearing members in the clearing of securities transactions (for details cf. Sect. 8.2.3).

In summary, the clearing of transactions by the CCP mainly comprises services in connection with the conclusion, collateralisation and settlement of

[14] Novation means the substitution of a new contract for an old one. The new agreement extinguishes the rights and obligations that were in effect under the old agreement.

transactions. There are benefits for market participants as well as for the stability of the financial markets in the (central) clearing of transactions through a CCP, compared to the bilateral settlement of transaction.

- The clearing process involves netting a set of transactions to obtain one settlement figure, so it increases settlement efficiency.
- The CCP can provide post-trade anonymity because (at least in the open offer concept) the clearing members see the CCP only as their counterparty.
- Central clearing ensures that each clearing member has its fulfilment risk covered by the CCP. Therefore, a clearing member default will be handled by the CCP. And it will not, as is the case for bilateral settlement, impact a potentially large number of other market participants. Regulation requires CCPs for that purpose to have sufficient default management procedures in place.[15]
- Clearing members (and non-clearing members) are able to achieve significant capital efficiencies for their transactions with a CCP. That's because regulation takes into account a reduced fulfilment risk for such transactions in the context of capital requirements for credit institutions. Transactions of clearing members that are cleared through a CCP are considered with a lower risk weight. The same applies to transactions of non-clearing members on the condition that certain segregation requirements[16] are met (Fig. 8.3).[17]

8.2.3 Admission of Clearing Members

The admission requirements of clearing members are outlined in the rules (clearing conditions) of the CCP. These rules are required under regulation to be transparent and to allow for the non-discriminatory access to the CCP. At the same time, they are also required to take into account the risk associated with the

---

[15] Art. 48 of the Regulation (EU) No. 648 2012 of the European Parliament and of the Council of 4 July 2012 on OTC derivatives, central counterparties and trade repositories (EMIR) requires CCPs to have detailed procedures in place to be followed in case of default of a clearing member. A CCP is required to take prompt action to contain losses and liquidity pressures resulting from defaults and shall ensure that the closing out of any clearing member's positions does not disrupt its operations or expose the non-defaulting clearing members to losses that they cannot anticipate or control. Where a CCP considers that a clearing member will be in default, it shall promptly inform the competent authority before the default procedure is declared or triggered. A CCP shall verify that its default procedures are enforceable. It shall take all reasonable steps to ensure that it has the legal powers to liquidate the proprietary positions of the defaulting clearing member and to transfer or liquidate the clients' positions of the defaulting clearing member.

[16] These requirements include individual client segregation and omnibus client segregation as further explained in Sect. 8.2.4.2 of this chapter.

[17] Own fund requirements of clearing members and their clients for exposures to a CCP are provided for in detail in Art. 300 et seq. of the Regulation (EU) No. 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending Regulation (EU) No. 648/2012.

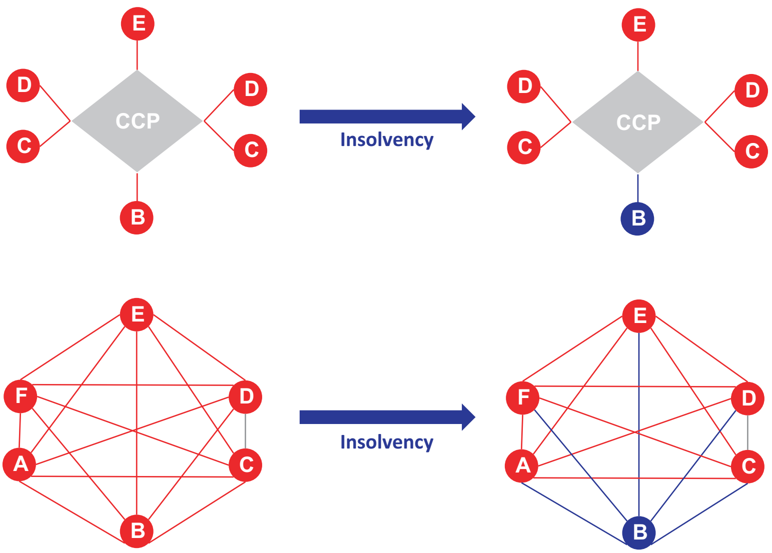


Fig. 8.3 Bilateral vs. central clearing

admission of clearing members.[18] On the one hand, admission requirements should not unreasonably prevent companies from access to clearing services. This is even more important for financial instruments such as OTC derivatives. Here regulation has imposed or will impose on market participants the obligation to clear transactions through a CCP (clearing obligation).[19] On the other hand, admission requirements for clearing members must be seen as a first step in a comprehensive system of safeguards designed by the CCP to control the risk of a clearing member failing. As a result, only companies that meet certain stringent criteria are eligible for clearing membership. The main admission requirements are as follows:

- Appropriate regulatory oversight as determined by the CCP must be demonstrated by companies applying for clearing membership at the CCP. Furthermore, they must be authorised and entitled by their respective regulatory authorities to operate in the custody and loan business as well as to accept receipt of margins in the form of cash and/or securities. In certain cases, e.g. for state organisations or clearing members that are only entitled

[18] Art. 37 (1) EMIR requires CCPs to establish, where relevant per type of product cleared, the categories of admissible clearing members and the admission criteria. Such criteria shall be non-discriminatory, transparent and objective so as to ensure fair and open access to the CCP and shall ensure that clearing members have sufficient financial resources and operational capacity to meet the obligations arising from participation in a CCP. Criteria that restrict access shall be permitted only to the extent that their objective is to control the risk for the CCP.

[19] The clearing obligation procedure for OTC derivatives transactions is set out in detail in Art. 4–6 EMIR.

to clear specific transactions, CCPs in their rules may waive the requirement of regulatory oversight of clearing members.

- Applicants must show evidence of a certain minimum level of liable capital[20] and must pay a contribution to the CCP default fund,[21] accessed by the CCP in the event of a clearing member default. In general, both the level of liable capital and the level of contribution to the clearing fund depend on the status of the clearing member (general clearing member or direct clearing member), as well as upon the markets that each clearing member clears.
- Additional admission requirements include different accounts that clearing members need to establish. Every clearing member must hold accounts at a CSD for the delivery of securities, the deposit of collateral and the execution of settlement-related payments. Furthermore, CCPs require evidence of cash accounts in the respective trading currency for clearing-related payments. These cash accounts, wherever available, should be held with central banks in order to reduce settlement bank risk.
- To participate in the clearing of transactions, not surprisingly, it is further required that clearing members have technical connectivity to the clearing system of the CCP and qualified back-office staff. A sufficient qualification of back-office staff is assumed if the test for clearing staff members offered by the CCP has been passed successfully.

8.2.4 Legal Relationships to Clearing Members and Non-clearing Members

Legal relationships are established by the CCP with clearing members and non-clearing members, in clearing transactions executed on the exchange.

8.2.4.1 Bilateral Clearing Agreements with Clearing Members

The CCP's clearing services are conducted on the basis of standard agreements that are entered into between the CCP and each single clearing member. These agreements set out the terms and conditions between the CCP and the clearing member

[20] Liable capital means available own funds of the applicant pursuant to the Regulation (EU) No. 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending Regulation (EU) No. 648/2012 in an amount determined by the CCP. Applicants not subject to the own fund requirements under such regulation must have available equivalent regulatory capital. Regulatory capital is considered equivalent when it is (a) used as a measure of adequate solvency for the applicant by its competent supervisory authority, (b) reported to the applicant's competent supervisory authority on a regular basis and (c) audited at least yearly.

[21] Pursuant to Art. 42 (1) EMIR, a CCP shall maintain a pre-funded default fund to cover losses that exceed the losses to be covered by margin payments of the clearing members, arising from the default, including the opening of an insolvency procedure, of one or more clearing members.

with the clearing of the clearing member's own securities transactions and securities transactions on behalf of clients of the clearing member. These clients may be non-clearing members (i.e. clients admitted to trading on the exchange), or clients that are no trading participants. In the latter case, the clearing members are at the same time trading participants. Transactions on behalf of their clients result from the matching of orders or quotes that have been entered by the clearing members into the trading system of the exchange.

The clearing agreements between the CCP and the clearing members contain, inter alia, rules on:

- The legal relationship between both parties in general
- The provision of margin collateral in the form of securities or cash for own transactions, and customer-related transactions
- Cash clearing and clearing currency
- Required debit instruction for the processing of cash payments and authorisation of the CCP (via power of attorney) to provide delivery instructions
- Representations of the clearing member and the CCP, e.g. regarding their capacity to enter into the agreement and existence of all required regulatory licenses and creditworthiness (no moratorium, opening of insolvency proceedings, etc.)
- Governing law and jurisdiction
- Amendments to the agreement that may be made unilaterally by the CCP
- Termination of the agreements

Furthermore, the clearing agreements lay out the rights and obligations of the CCP and the clearing members on the specific clearing models offered by the CCP, a relevant detail for clearing transactions by clearing member on behalf of non-clearing members. We will turn to these clearing models in more detail under (b) below.

The clearing agreements bring together the rules, or clearing conditions of the CCP that are acknowledged by the clearing members in their signing of the agreements. The clearing conditions contain all the rules for providing the clearing services, in particular the following:

- Scope of the clearing services
- Admission of clearing members and termination of the clearing membership by the CCP or the clearing member
- Delivery of securities and payment of the purchase price to settle securities transactions executed on the exchange
- Netting procedure
- Provision, valuation and enforcement of margin in the form of securities or cash
- Contributions to the clearing funds
- Procedure in case of failure of a clearing member to deliver securities (buy-in or cash settlement)
- Procedure in the event of a default of a clearing member (other than buy-in or cash settlement in case of failure to deliver securities) or of the CCP

- Special provisions for the different clearing models offered by the CCP, in particular on the provision of margin, segregation of assets and positions, accounts and event of default
- Confidentiality obligations of the CCP
- Governing law and place of jurisdiction
- Unilateral amendments to the clearing conditions made by the CCP

Clearing members pay for the services of the CCP, so the clearing agreements also include the CCP's price list setting out the fees for the different services.

8.2.4.2 Tripartite Clearing Agreements with Clearing Members and Non-clearing Members

The CCP, a clearing member and a non-clearing member may enter into a standard tripartite clearing agreement to clear securities transactions on behalf of non-clearing members, in addition to the clearing agreements between the CCP and clearing members. Clearing of transactions of non-clearing members involves additional complexity due to the requirement to segregate positions and assets of clearing members and non-clearing members. These tripartite clearing agreements provide for terms and conditions that apply between:

- The CCP, the clearing member and the non-clearing member
- The CCP and the clearing member, on the one hand, and between the clearing member and the non-clearing member, on the other

The tripartite clearing agreements, like bilateral clearing agreements, bring together the clearing conditions and the price list of the CCP. Provisions already outlined above (a) for bilateral agreements are included.

The tripartite agreements in particular seek to protect non-clearing members in the event of a clearing member default. Following high-profile insolvencies, regulatory reform and increasing client demand, CCPs have introduced specific clearing models. These require the segregation of the clearing members' positions (transactions) and assets (collateral) from those of their non-clearing members. In the event of a clearing member default, the clearing models allow for the transfer of the positions and assets of a defaulting clearing member's clients to a solvent clearing member. In some instances, this may mean the orderly liquidation of the clients' positions and the return of excess collateral to the clients. In general, two different clearing models can be distinguished, the individual client segregation and the omnibus client segregation. The actual level of protection depends on the level of segregation of positions and assets that clients (non-clearing members) select.

Accordingly, the CCP, a clearing member and a non-clearing member may enter into a standard tripartite clearing agreement either under the individual client segregation model or under the omnibus client segregation model.

- In the case of individual client segregation, the CCP is obliged under the agreement to segregate through individual accounts per non-clearing member

positions and assets (i.e. securities transactions and collateral) of the individual non-clearing member and of its clearing member. Under this agreement, if the clearing member defaults, the non-clearing member can choose to transfer its positions and assets to a new clearing member. In cases where the transfer is successful, the non-clearing member's segregated positions and assets will be transferred to the new clearing member without the kind of liquidation risk inherent in clearing models with no segregation.
- Non-clearing members who choose the omnibus client segregation accept a lower level of protection in case their clearing members default. In this model, the tripartite clearing agreements do not require segregation per individual non-clearing member. Rather, only the positions and assets of all non-clearing members together on behalf of which the clearing member clears securities transactions are segregated from the positions and assets of the clearing member. Therefore, in a clearing member default, non-clearing members are not able to decide individually on the transfer of their specific positions and assets to a new clearing member. Rather, only a transfer is possible of the positions and assets of all the non-clearing members.[22]

8.2.4.3 Further Agreements

In addition to the clearing agreements mentioned above, further agreements are entered into by the CCP, clearing members and non-clearing members. Clearing members and non-clearing members must be granted access by the CCP to the technical clearing system to participate in the clearing of securities transactions. This access and the use of the clearing system are based on technical connection agreements executed by the CCP with clearing members and non-clearing members. These agreements under the General Terms and Conditions provide for:

- Details of the technical connection, such as access alternatives, interfaces and specifications
- Liability of the parties, in particular of the CCP, in case of disruption or default of the clearing system
- Unilateral amendments to the agreements made by the CCP
- Confidentiality obligations
- Governing law and place of jurisdiction
- The fees for the technical connection with a price list included into the agreement

[22] Pursuant to Art. 39 (1) and (2) EMIR, the CCP (a) shall offer to keep separate records and accounts enabling each clearing member to distinguish in accounts with the CCP the assets and positions of that clearing member from those held for the accounts of its clients (omnibus client segregation) and (b) shall offer to keep separate records and accounts enabling each clearing member to distinguish in accounts with the CCP the assets and positions held for the account of a client from those held for the account of other clients (individual client segregation).

Regulation and/or the rules of the CCP may provide for committees to be established for the purpose of advising and assisting the CCP on certain aspects of the clearing procedure. For example, a default management committee may be established for the default of a clearing member and the subsequent default management process. Participation in this committee is based on standard agreements between the CCP and clearing members. This defines the rules for the participation of the clearing member's employees in the default management committee. These agreements, most notably, include rules on the confidential treatment of data provided during the default management process.

8.3 The Custody and Settlement Layer

We finally turn to the custody of securities and the settlement of securities transactions and consider the roles of the CSD and its customers (Sect. 8.3.1) as well as the legal relationships between them (Sect. 8.3.2).

8.3.1 CSDs and Their Customers

The CSD provides custody and settlement services to various groups of customers. The CSD handles the holding and administration of securities, in particular, collecting dividends, withholding taxes and accounting of stock dividends and splits. For the settlement of securities transactions, the CSD services are relevant for trading participants, intermediary banks, clearing members and CCPs with legal relationships established by the CSD.

- Trading participants have to conduct the settlement of securities transactions through a CSD recognised by the exchange and through an accounting relationship with a central bank, or other payment bank. This requirement can be met either by a direct legal relationship of the trading participant with the CSD or through an intermediary bank that maintains accounts, and a corresponding legal relationship with the CSD.
- As already set out (Sect. 1.4.2.3), for securities transactions cleared through a CCP, the rules of the CCP require clearing members to hold accounts at a CSD for the settlement or transactions, and the deposit of collateral.
- Furthermore, if collateral in the form of securities has to be deposited by clearing members in an account of the CCP, the CCP requires access to a CSD for the collateralisation of securities transactions.

Settlement services provided by the CSD ensure delivery versus payment of the securities traded on the exchange, and payment of the purchase price within a predefined period of working days. Thereby, a simultaneous movement of cash and securities in opposite directions takes place between the parties to the transactions, i.e. for securities transactions cleared through a CCP between the CCP and its clearing members. Settlement is usually due in 2 or 3 working days after conclusion of a securities transaction. A physical exchange of cash and products rarely takes place anymore, as the securities exchanged are all standardised and dematerialised. Therefore, book entries in an accounting system have replaced paper certificates as evidence of ownership. CSDs deploy technical settlement systems that provide efficient platforms for the entry and processing of instructions for the settlement process. In case of securities held abroad with a foreign CSD, the CSD establishes a link and legal relationships, either directly with the foreign CSD or indirectly through a correspondent bank. For transactions cleared through a CCP, after netting of transactions by the CCP, the settlement is performed by the CSD only of the balance of purchases and sales for each security (Fig. 8.4).

8.3.2 *Legal Relationships to Customers*

The legal relationship between the CSD and its customers—trading participants, intermediary banks, CCPs and clearing members—usually results from the customers' signing of the CSD's account-opening forms. This signifies that the customers accept the CSD's General Terms and Conditions.

The General Terms and Conditions set forth the rules governing the provision of services and products by the CSD to its customers. Basically, they apply to the

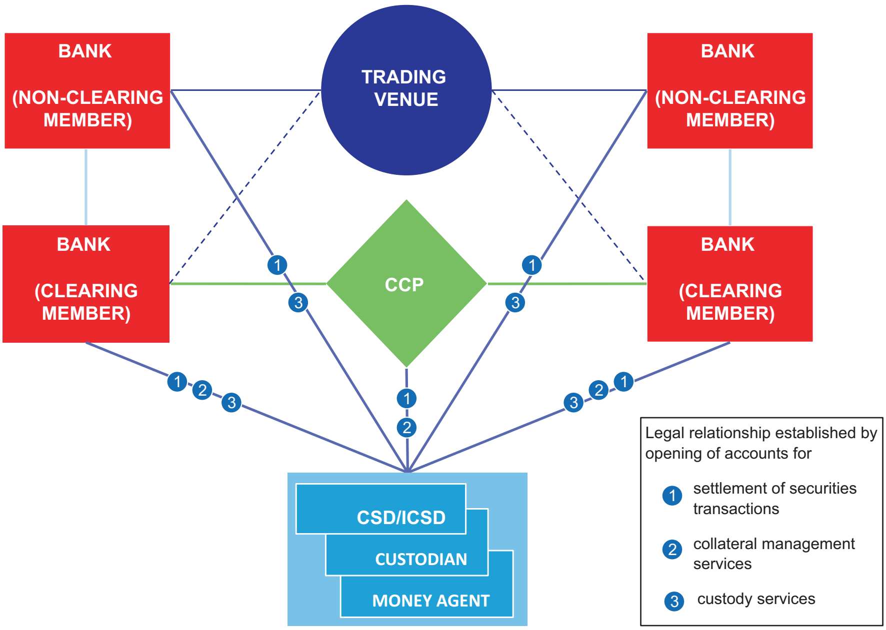


Fig. 8.4 Legal relationships on the settlement and custody layer

provision of all services by the CSD to its customers and to all accounts of the customer with the CSD. The main provisions of the General Terms and Conditions include:

- The admission and deposit of securities to collective safe custody and their withdrawal
- Customers' general obligations relating to settlement of instructions
- The administration of securities in custody with the CSD (that is, securities account statement, disclosure of information, redemption of securities, coupon renewal, principal maturities, corporate actions, services relating to general meetings, tax-related services)
- The order for transactions of securities, and the transfer of securities held in collective safe custody and in non-collective safe custody
- Collective safe custody of registered shares
- Settlement finality rules[23]
- Liability regime
- Applicable fees

In addition to the general terms and conditions, special conditions usually prevail over the CSD's General Terms and Conditions in the case of particular services by the CSD to its customers.

More specifically, CSDs offer collateral management services as a collateral agent for two customers acting, respectively, as collateral receiver and the collateral giver. These services may be governed by special conditions, and include collateral eligibility check, collateral and exposure valuation, margin calls, substitution and reporting.

[23] In the settlement of securities transactions, a distinction must be made between (a) the finality of instructions to transfer securities and/or cash amounts in a clearing and settlement system in the meaning of Directive 98/26/EC of the European Parliament and of the Council of 19 May 1998 on settlement finality in payment and securities settlement systems (EC Directive on Settlement Finality) and (b) the lawful settlement of securities and/or cash delivery obligations. Finality means the legal enforceability and irreversibility of such a transfer instruction, as defined by the rules of the CSD, once it has been entered into such a system, provided that there are sufficient positions of cash and securities. If the customer who enters a transfer instruction is insolvent, "finality" affords the other customers and the processing system itself protection against the unwinding risk, that is, the reversal of the cash and securities transactions, which are in the process of being settled. From a legal point of view, the settlement of delivery obligations takes place when the counterparties to a securities transaction have performed their entire obligations (such as the transfer of rights to a security or effecting a cash payment) so that all claims arising from the transaction have been satisfied in full and are consequently extinguished.

Chapter 9
Financial Market Regulation

Reto Francioni, James H. Freis Jr., and Alexandra Hachmeister

9.1 Setting the Stage

A well-functioning financial system is indispensable for economic growth. To that end, the overall objective of the current regulatory reform process is to guarantee, globally, financial market stability.[1] In this context, one of the most important preconditions for stability is a system of working financial market infrastructure (FMIs). Every analysis of financial markets must naturally account for its infrastructure regulations.

This need came to the fore most clearly after the last financial crisis. The central importance of regulatory reform, more than ever before, defined the strategy of financial market operators and participants.

The twentieth century saw a multitude of economic shocks connected to market crashes, along with cross-border crises as the ripple effects. This begs the following question: What turning points in history help explain this phenomena? A good place to start is the opening of national financial market to the global stage. This occurred with the collapse of the Bretton Woods system[2] in the early 1970s in tandem with advances

---

[1] Financial market stability is a multidimensional construct with the most important factors being system stability and investor/customer protection and the core element of guaranteeing fair, safe, and sound markets through the elimination of risks for market infrastructure and participants that might also affect the internal market and the real economy.

[2] A comprehensive analysis regarding the development and impact of the Bretton Woods system and its key performers can be found in Benn Steil, 2013: "The Battle of Bretton Woods," Princeton University Press.

R. Francioni (✉)
Wirtschaftswissenschaftliche Fakultät, University of Basel, Basel, Switzerland
e-mail: reto.francioni@unibas.ch

J.H. Freis Jr. • A. Hachmeister
Deutsche Börse AG, Mergenthalerallee 61, 65760 Eschborn, Germany
e-mail: James.freis@deutsche-boerse.com; Alexandra.hachmeister@deutsche-boerse.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_9

in the electronization of the financial system. The financial globalization resulted in a liberalization of international capital flows and access to capital with better opportunity for risk diversification. What's more, the Digital Revolution that facilitated a significant improvement in cumulated trading volume, latency, and overall performance raised the stakes, adding potential new triggers for future crises. It also included an increased vulnerability—in scope, magnitude, and time—to changes in the financial systems in other parts of the world with the threat of financial **contagion**. This was exhibited in the bursting of the US housing bubble in 2007 and the global financial crisis that followed.

The mechanism of those crises (cf. Fig. 9.1) can often be defined by a specific pattern. American economist Hyman Minsky in his *Financial Instability Hypothesis*[3] termed *procyclicality of the financial system*, and defined it as "*mutually reinforcing mechanisms between the performance of the financial system and the economy, through which the financial system can amplify fluctuations and possibly cause or exacerbate instability.*"[4]

Periods of severe financial market distress are usually followed by intensified surveillance and regulatory processes. That is accompanied by rather conservative, low-risk investment practices resulting in balanced interest rate-loan proportions. As the rise of credit and asset prices mirrors market growth, the willingness to take more risk rises accordingly. That can happen on the "official" side through deregulation—as in the late 1990s, especially in the USA, with the far-reaching repeal of the historical divisions between investment and commercial banking under the Depression-era **Glass-Steagall legislation**. On the participant side, it can happen with high-risk products and investment practices. The speculative phase shows the introduction of more complicated financial instruments (see **financial innovation—securitization**, cf. Fig. 9.2) as well as high-speculative investment practices to help secure future profits. As risk increases, the financial system usually fails to build sufficient capital and liquidity buffers to cover defaults. This brings changes in the risk structure of financial market as well as in the risk-return relation. And it is

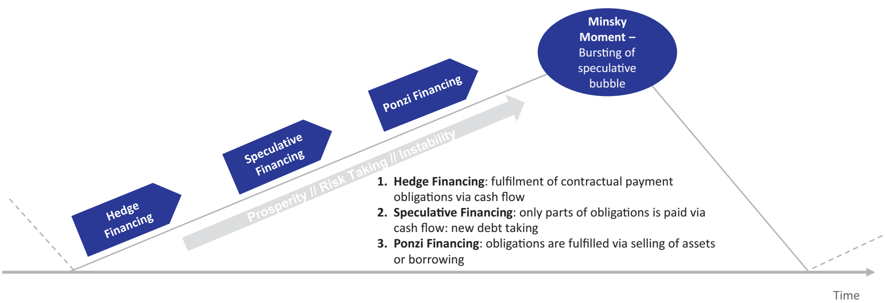


Fig. 9.1 Procyclicality of the financial system (based on [17])

[3] See [17].

[4] FSF/[2], p. 1.

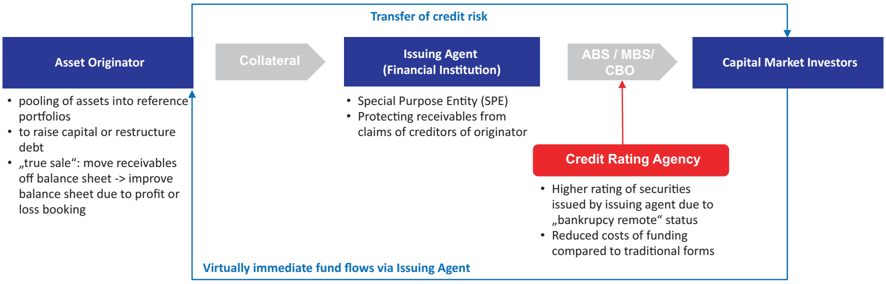


Fig. 9.2 Securitization process (based on [15])

amplified in the final stage when financial institutions can no longer mitigate losses.[5]

The consequences are unstable financial structures that are increasingly susceptible to economic change combined with a general inability to accurately monitor, evaluate, and validate risk. The shock of an eventual bursting of the speculative bubble has notable macroeconomic effects and vice versa—i.e., financial (market) crises—due to the ever-growing (global) interconnectedness of financial markets (over time and over the economic system).

These market interlinkages—cross-sector and cross-border—present significant **systemic risk** to the overall stability of the economic system. The **Bank for International Settlements** (BIS) and FSB define systemic risk as "*a risk of disruption to financial services that is caused by an impairment of all or parts of the financial system and has the potential to have serious negative consequences for the real economy*."[6] Additionally, a market, instrument, or institution is considered **systemically important** "*if its malfunction causes widespread distress, either as a direct impact or as a trigger of broad contagion.*"[7]

Take the current financial crisis as an example. World financial assets have grown much faster than the average gross domestic product between 1980 and 2007—the time of the credit boom.[8] At the same time, this was accompanied by continuous deregulation that facilitated a change from the classic business model of banks to an **originate-and-distribute model (shadow banking system[9])**, a model characterized by loans that were no longer required to be included in the bank's balance sheets. Loans could instead be pooled into a reference portfolio/asset pool and sold

[5] For an analysis of procyclicality effects leading to the financial crisis of 2009 refer to [10].

[6] http://www.bis.org/publ/othp07.pdf: [14], p. 5. What is described here relates also to the change in the overall market risk structure that has to be adequately covered by regulation and surveillance as well as, on participant's side, by monitoring and operations.

[7] http://www.bis.org/publ/othp07.pdf: [14], p. 5.

[8] See [9]: https://hbr.org/2008/09/new-thinking-for-a-new-financial-order.

[9] Similar to banks, shadow banks perform term transformations by financing long-term assets with short-term loans (breaking the "golden rule" of duration symmetry), outside the banking system and without access to central bank liquidity. Basic capital requirements that are valid for banks can thus be avoided.

(securitization, cf. Fig. 9.2).[10] The defective pricing of risk that stemmed from those products (especially **collateralized debt obligations (CDO)**, i.e., **asset (ABS)-** or **mortgage (MBS)-backed securities** as the focus of speculative investment practices) contributed significantly to the instability of the system.[11]

CDOs allowed subprime mortgages to be originated and sold to investors worldwide. At the same time rating agencies, using complex risk models to calculate the risks of these synthesized financial products, have wrongly assessed their default probabilities as well as their underlying individual mortgages. As these practices evolved, even lower rated tranches from various MBS CDOs were repackaged into new CDOs with a higher rating. In the subsequent downturn, this ultimately proved illusory. The high ratings that resulted fostered international demand and investment flow into the US housing bubble.

At some point, interest could no longer be paid. The consequence was unexpected defaults. A consecutive striving for the liquidation of assets and a wish for cash by investors in return added to the downturn of the overall market. Banks, unable to compensate for losses due to an insufficient or unsecured capital basis—the capital base pressured downwards as part of the overall deregulatory trend—started to withhold short-term credit and to cut loans for the nonfinancial industry.

The market for short-term commercial paper dried up. This had negative effects on the economy and the labor market, spilling over beyond those that relied directly on this liquid money flow. Due to mutually reinforcing connections between the financial market and the economy, the events in the USA triggered losses on international stock and FX markets. That eventually led to a slowing global economy, decreased trading and tightened credit availability. In the aftermath of the US banking crisis, the recession was worsened in the EU by a Sovereign Debt Crisis. European States, unable to pay off interests, were forced to increase debt to fund bank bailouts and countercyclical fiscal measures.

Figure 9.3 depicts a schematic overview of this evolution of financial markets regulation in Europe. Shortly before the crisis, MiFID I and the “Financial Services Action Plan” started the harmonization process of the European financial markets landscape. Yet the focus was on self-regulation conducted by each financial institution individually rather than on centralized regulation by international supervisory bodies. Then, as a reaction to the events that led up to the global financial crisis, the **G20** states at the 2009 Pittsburgh Summit committed to a thorough reformation of surveillance structures and regulatory processes to help reinstate a sustainable stability of financial market infrastructures. They called this the new order of the financial industry.[12]

---

[10] Descriptive illustration of the securitization process provided by the International Monetary Fund in Jobst, 2008: http://www.imf.org/external/pubs/ft/fandd/2008/09/pdf/basics.pdf.

[11] Financial innovation, however, is not by itself an impairing factor for financial markets—on the contrary. The problem lies—as usual—in the specific construction, use, and rating of those instruments. For a detailed analysis of and outlook on the securitization market see [3].

[12] The term “new (financial) order” was coined in 2003 by Yale Professor and Nobel Laureate Robert J. Shiller: “The New Financial Order proposes a radically new risk management infrastruc-

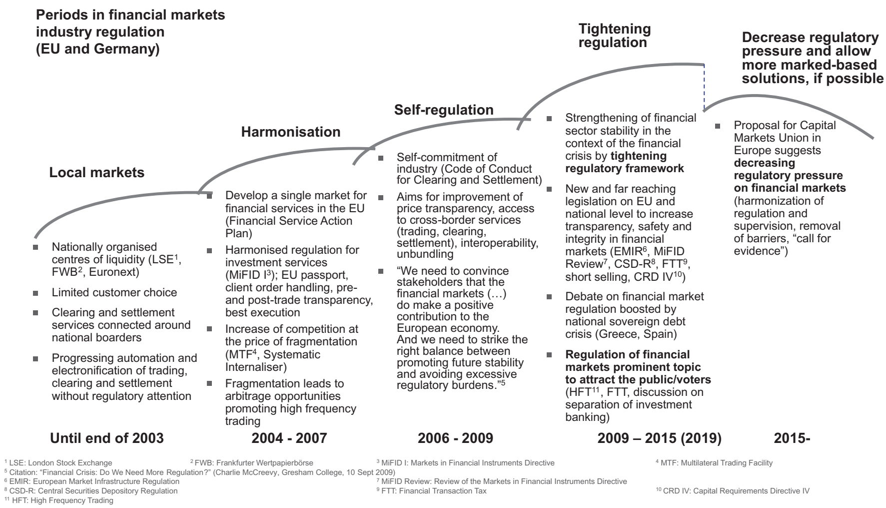


Fig. 9.3 Periods in European Financial Markets Regulation (for abbreviations see abbreviation list)

And so the concept of self-regulation was discarded and replaced by G20 principles for financial market regulation. As a result, the USA developed the Dodd Frank Wall Street Reform Act and the EU introduced maximum harmonization rules for trading venues and FMIs, namely EMIR, MiFID II, and CRD-IV. These structural reforms globally redesigned financial markets and impacted the market operating business today.

9.2 Global Financial Market Supervisory Structure and Regulatory Processes

At its core, the global financial market infrastructure can be considered an outcome of institutional interplay and personal relationships rather than as an isolated development that has followed political enforcement.[13] The first wave of modern—that is to say, global—financial market regulation started post-World War II with the Bretton Woods system of international exchange rate standards. Bretton Woods saw the establishment of the **International Monetary Fund** (IMF), the **World Bank** (WB), and the **World Trade Organization** (WTO)—international organizations that continue to influence and shape financial markets worldwide.

ture to help secure the wealth of nations: to preserve the billions of minor—and not so minor—economic gains that sustain people around the world" ([18], p. IX).

[13] Blair et al., 2012, p. 473. [6], p. 2.

However, with the increasing internationalization of capital mobility, and following the interconnectedness of financial markets, the Bretton Woods system, as outlined earlier in this chapter, basically lost its scope of responsibility. Therefore, its *raison d'être* was ultimately extinguished. What was defined by treaties before was now increasingly defined by the market itself where fixed exchange rates were no longer required[14]: "*National authorities responsible for monetary and financial politics, such as central banks and regulatory authorities assumed a more important role. Their action shaped the evolution of an increasingly integrated international financial system.*"

The reestablishment of a global system of supervision and regulation in the wake of the recent financial crisis enhanced this process of financial market integration. Supervision today presents itself as a cascade, starting at the international level and building a thorough net that passes through to each individual jurisdiction (cf. Fig. 9.4).[15]

As a result of the Pittsburgh summit, the G20 agreed on, and committed themselves to, among others, the *principles for financial market infrastructure*.[16] Those principles were tied to the idea that nearly every market-specific regulation is linked to a global minimum standard seeking to ensure a common basis for financial market regulation worldwide: "*In general, these standards are expressed as broad principles in recognition of FMIs' differing organisations, functions, and designs, and the different ways to achieve a particular result. In some cases, the principles also incorporate a specific minimum requirement (such as in the credit, liquidity, and general business risk principles) to ensure a common base level of risk management across FMIs and countries.*"[17]

Because of those developments, nothing has had such a profound influence on the business strategies of financial service providers and market participants as global regulation and supervision. The former describes the monitoring of the behavior of financial market participants and the enforcement of legislation; the latter is defined as the process of rulemaking and legislation underlying the supervisory framework (Fig. 9.4).

9.2.1 Global Supervisory Structure

The **Financial Stability Forum (FSF)** was founded in 1999 on the initiative of the then Bundesbank Chief, Hans Tietmeyer, in response to the shortcomings in regulation and oversight. In particular, this became apparent in cross-border cooperation in the supervision of increasingly international financial institutions and markets

[14] Blair et al., [13], p. 475.

[15] The whole G20 Pittsburgh Leader's Summit is available at [11]: http://www.bundesregierung.de/Content/DE/StatischeSeiten/Breg/G8G20/Anlagen/G20-erklaerung-pittsburgh-2009-en.pdf?__blob=publicationFile&v=2.

[16] The list of principles is available online: http://www.bis.org/cpmi/publ/d101a.pdf.

[17] See [5], p. 5.

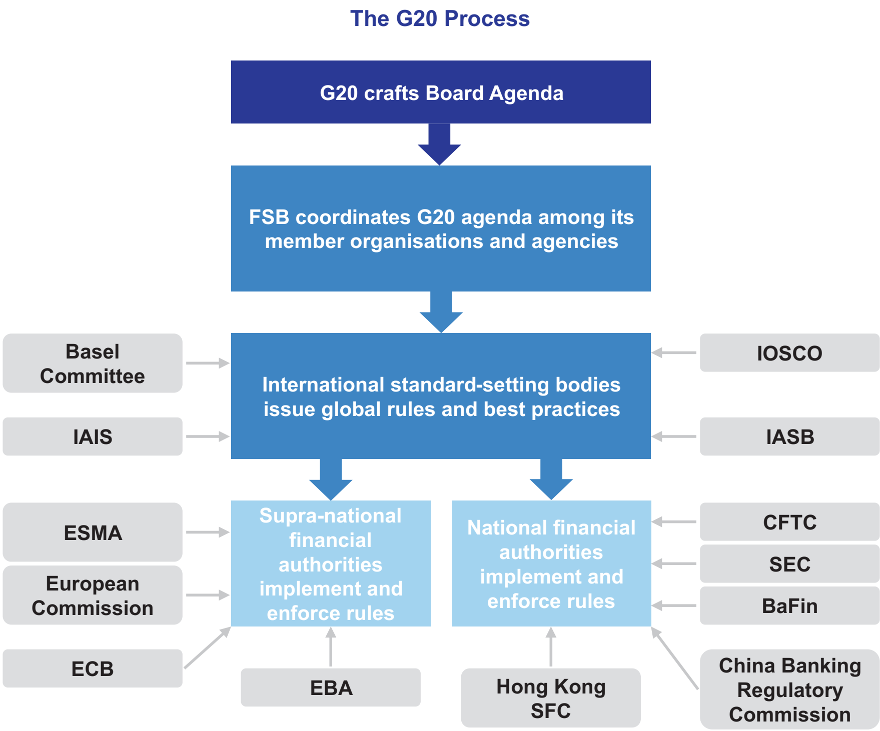


Fig. 9.4 G20 Supervisory Cascade (based on the Atlantic Council of the United States [1]), p. 12 (please refer to abbreviation list for details))

following the Asian financial crisis. As the financial crisis of 2007–2009 once again revealed massive shortcomings in regulation, the FSF would later expand in influence and importance through the G20. In 2009, it was renamed as the **Financial Stability Board (FSB)**, and had its mandate broadened to include all G20 jurisdictions. Today, it functions as the central organization that oversees the implementation of G20 principles in all member states through their member institutions and organizations. In this function, it is supported by International **Standard-Setting Bodies (SSBs)** that develop and issue concrete regulatory reforms and respective best practice advice according to the G20 agenda.

9.2.1.1 The Group of 20 (G20)

The G20 defines itself as the global forum for international economic issues, comprising the heads of state of the world's leading developed and emerging economies, as well as the European Union. According to its website, in 2014 the G20 represented more than two-thirds of the world's population, over 80 % of the international economic power, and approximately 75 % of the overall global trade.

The G20, founded in 1999, has now evolved to become the leading body in the world in setting international principles for financial markets regulation. The organization (among others) is dedicated to economic growth, financial supervision, and international market regulation The financial markets issues were originally addressed at more of a technical level through meetings of finance ministers and central bank governors.

But it has since increasingly assumed more political prominence. In response to the global financial crisis, leader's summits have been hosted at least once a year since 2008 with an annually rotating presidency to develop the G20 agenda and to oversee its implementation within respective member states.[18]

According to its official G20 Pittsburgh Summit statement,[19] leaders agreed:

- *To launch a framework that lays out the policies and the way we act together to generate strong, sustainable and balanced global growth.*
- *To make sure our regulatory system for banks and other financial firms reins in the excesses that led to the crisis.*
- *To reform the global architecture to meet the needs of the twenty-first century.*

9.2.1.2 The Financial Stability Board

The Financial Stability Board (FSB) coordinates national authorities responsible for financial stability, international financial institutions and SSBs, sector-specific international groups of regulators and supervisors, and committees of central bank experts. The FSB's objective is to identify systemic risk on a timely basis and to develop policies based on the G20 agenda that supports risk containment and prevention. The FSB's work is aimed at ensuring the development of efficient financial policies with coherent national implementation. The upshot is a level playing field of international market regulation across sectors and jurisdictions.

The FSB plenary is the sole decision making body of the FSB, consisting of central bank governors, heads of the main supervisory/regulatory agencies, and deputy finance ministers as well as the chairs of the main SSBs and other international and European institutions. On the working level, it is supported by a steering committee as well as three standing committees for support in the specification of legislations. These three standing committees are for assessment of vulnerabilities, supervisory and regulatory cooperation, and standards implementation. Nonetheless, the FSB has no legislative power, so its decisions are not binding. Members are required to commit themselves to adopting international standards and to implementing them into national legislation.[20]

[18] www.g20.org.

[19] The whole G20 Pittsburgh Leader's Summit is available at Available online: http://www.bundesregierung.de/Content/DE/StatischeSeiten/Breg/G8G20/Anlagen/G20-erklaerung-pittsburgh-2009-en.pdf?__blob=publicationFile&v=2.

[20] Further information available at http://www.financialstabilityboard.org/.

9.2.1.3 International Standard Setting Bodies[21]

The oldest group of national financial regulatory and supervisory authorities is the **Basel Committee on Banking Supervision (BCBS)**, established in 1974 by the governors of the G-10 central banks[22] in the wake of the failure of Herstatt Bank of Cologne, Germany; the failure had systemic repercussions. The national regulatory and supervisory authorities for the securities and the insurance sectors subsequently established the **International Organization of Securities Commissions (IOSCO)** and the **International Association of Insurance Supervisors (IAIS)**, respectively, albeit in each of these cases with a view towards universal, global membership.

- The BCBS provides a forum for regular co-operation among member countries on banking supervisory matters. The objective is to enhance understanding of key supervisory issues and to improve the quality of banking supervision worldwide. BCBS' first major document in 1975 was the Basel Concordat, since renamed Principles for the Supervision of Banks' Foreign Establishments. This document provides guidance on the sharing of supervisory responsibilities among home and host country supervisors. BCBS has since become most broadly known for having established rules concerning minimum capital requirements for internationally active banks. First agreed in 1988, it is now adopted globally in its third iteration as **Basel III**. A fourth round is under development. In more recent years, the focus of the SSBs has shifted away from developing new standards, and more towards deepening and ensuring consistent implementation (in particular through aspects of peer reviews). Under this arrangement, experts from multiple jurisdictions review the status of implementation in the individual member states. This is exemplified by the BCBS' establishment of the **Regulatory Consistency Assessment Programme (RCAP)** in 2012 to facilitate the consistent implementation of the G20 agenda in banking regulation, primarily with respect to capital and liquidity requirements under **Basel III**. The RCAP subsumes two steps:
  - Monitoring of implementation
  - Assessing the consistency among member jurisdictions to secure that minimum standards are met, possible regulatory gaps identified, and banking institutions are performed accordingly[23]

[21] For a more detailed review of the history of the development of the international financial standard setting bodies, see Mario Giovanoli, "A New Architecture for the Global Financial Market: Legal Aspects of International Financial Standard Setting," in Mario Giovanoli, ed., International Monetary Law: Issues for the New Millennium (Oxford University Press: Oxford, 2000).

[22] The G-10 central banks were for decades the institutions which governed the Bank for International Settlements (BIS) in Basel, Switzerland, which hosts the secretariats of a number of the international standard setters including the FSB, BCBS, and IAIS. The G-10 consisted of Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Sweden, the USA, and the UK; Switzerland was later added as its 11th member. The G-10 groupings have more recently evolved largely to reflect broader G-20 participation.

[23] http://www.bis.org/bcbs/.

- IOSCO is the international policy forum for national regulators of securities and futures markets. This forum develops and promotes standards of securities regulation to maintain efficient and sound markets, and to establish standards for effective supervision of the international securities market.
- IOSCO, like the BCBS, established an assessment committee to oversee the implementation of IOSCO principles and standards within member jurisdictions. The AC conducts country reviews based on members' self-assessment to evaluate the overall status of implementation and provide guidance, thematic reviews to identify possible regulatory gaps in standards and principles regarding specific topics, and general maintenance work on principles and methodologies (in terms of understanding and timeliness of data).[24]
- IAIS represents insurance regulators and supervisors across borders and jurisdictions. IAIS' mission is to promote an effective and globally consistent regulation and supervision of the insurance industry in order to develop and maintain fair, safe, and stable insurance markets for the benefit and protection of policyholders, and to contribute to global financial stability.[25]

In 1996, BCBS, IOSCO, and IAIS began to coordinate their work across the three sectors through a so-called **Joint Forum.** In so doing, they paid special attention to international financial conglomerates that had activity crossing both sectoral and jurisdictional lines. According to its mandate:"*[t]he objective of the Joint Forum is to support banking, insurance and securities supervisors in meeting their regulatory and supervisory objectives and, more broadly, to contribute to the international regulatory agenda in particular where risks exist across or in gaps between the three supervised sectors.*"[26] The coordination is evidenced in the similar approach to the development and promulgation of "Core Principles" for the supervision of the respective banking, securities, and insurance sectors.

Two expert groups are represented in the FSB: the Committee on the Global Financial System, a central bank group monitoring financial stability issues (but not itself a standard setter), and the **Committee on Payments and Market Infrastructures (CPMI)**.

- CPMI, formerly known as the **Committee on Payment and Settlement Systems (CPSS)**, promotes the safety and efficiency of payment, clearing, settlement, and related arrangements. In this way, it supports financial stability within the global economy. CPSS also monitors and analyses developments in these arrangements, both within and across jurisdictions, serving as a forum for central bank cooperation in related oversight, policy, and operational matters, including the provision of central bank services.[27] One of the most important regulatory frameworks for exchange organizations was published by CPSS-IOSCO in 2012: The Principles for Financial Market Infrastructures.[28]

[24] http://www.iosco.org/about/?subsection=display_committee&cmtid=19.

[25] http://www.iaisweb.org/index.cfm?event=showHomePage&persistId=50D32488155D896B005D848D69E17DBA.

[26] http://www.bis.org/bcbs/jfmandate.html.

[27] http://www.bis.org/cpmi/index.htm.

[28] Available online: http://www.bis.org/cpmi/publ/d101a.pdf.

The other organizational members of the FSB are the IMF, the WB, the **Organisation for Economic Cooperation and Development (OECD)**, and the Financial Action Task Force (FATF). The FATF is the standard setter for the "market integrity" principles—the FATF Recommendations: International Standards on the Combating of Money Laundering and the Financing of Terrorism and Proliferation. Those issues have become a significant part of the financial regulatory agenda in the twenty-first century, as is clear with their inclusion as part of the G20 agendas. The G20 has long taken this crucial position: Within a jurisdiction's national framework, strong prudential financial standards are mutually reinforcing and strongly correlated with efforts to strengthen the integrity of financial markets' by combating money laundering, terrorist financing, corruption, and tax evasion.[29]

FSB Key Standards for Sound Financial Systems

Financial Regulation and Supervision

IAIS, Insurance Core Principles, Standards, Guidance and Assessment Methodology (2013)

BCBS, Core Principles for Effective Banking Supervision (2012)

IOSCO, Objectives and Principles of Securities Regulation (2010)

Institutional and Market Infrastructure

IOSCO, Principles for Financial Benchmarks (2013)

CPMI/IOSCO, Principles for Financial Market Infrastructures (2012)

FATF, FATF Recommendations on Combating Money Laundering and the Financing of Terrorism & Proliferation (2012)

WB, Insolvency and Creditor Rights Standard (2011)

OECD, Principles of Corporate Governance (2004)

International Accounting Standards Board (IASB), International Financial Reporting Standards (2002)

International Association of Deposit Insurers (IADI), IADI Core Principles for Effective Deposit Insurance Systems (2014)

International Auditing and Assurance Standards Board (IAASB), International Standards on Auditing (2014)

Macroeconomic Policy and Data Transparency

IMF, Code of Good Practices on Fiscal Transparency (2007)

IMF, Code of Good Practices on Transparency in Monetary and Financial Policies (2000)

IMF, General Data Dissemination System (1997)

IMF, Special Data Dissemination Standard (1996)

[29] See James H. Freis, Jr., "The G-20 Emphasis on Promoting Integrity in Financial Markets," in Mario Giovanoli and Diego Devos, eds., International Monetary and Financial Law: The Global Crisis (Oxford University Press: Oxford, 2010).

The FSB framework provides the foundation for a consistent approach at the international level to financial supervision. The FSB has compiled international financial standards and designated a number of key standards, including some mentioned above, that deserve priority implementation.[30] Standards and principles are either converted into national or European law, or have distinct influence on the respective processes within jurisdictions.

9.2.1.4 European System of Financial Supervision

The European Commission appointed Jacques de Larosière in 2009 to form a high-level expert group and establish a plan to "*repair*" the European system of regulation and supervision. The de Larosière report, with echoes of the G20/SSB process, proposed an approach that focused on the establishment of a **European System of Financial Supervision (ESFS)**.[31] In 2010, after ratification by the European Commission and Parliament, the system consisting of two individual supervisory authorities—a macro-prudential and a micro-prudential stream—was inaugurated (Fig. 9.5).[32] The ESFS aims in general at harmonizing the EU's supervision of financial markets via a standardized implementation mechanism of European Community Law. The Larosière report covers three different sectors of financial market regulation and supervision:

1. Regulation of instruments, services, and risk management in accordance with the G20 agenda (e.g., OTC derivative regulation, Basel III).
2. Supervisory and regulatory infrastructure in general: Establishing a harmonized single market (p. 27) with a consistent set of rules (p. 29)—"*Single Rulebook*"—and a minimum set of supervisory standards (p. 39) to establish a **level playing field** among EU member states.
3. European Supervisory Authorities (ESAs):

   (a) Strengthened role of the **European Central Bank (ECB)** regarding macro-prudential supervision (p. 42) in cooperation with the **European Systemic Risk Board** (**ESRB,** p. 46)
   (b) Establishment of new **ESAs** to manage micro-prudential supervision by defining common supervisory practices, developing technical standards, and coordinating a College of Supervisors in the EU (p. 48ff.)
   (c) Strengthened role of Competent Supervisory Authorities in the member states, where the supervision of domestic institutions remains (p. 52)

[30] See http://www.financialstabilityboard.org/what-we-do/about-the-compendium-of-standards/key_standards/.

[31] The complete report online: http://ec.europa.eu/internal_market/finances/docs/de_larosiere_report_en.pdf.

[32] De Larosière (2009:): 38: 46–48.

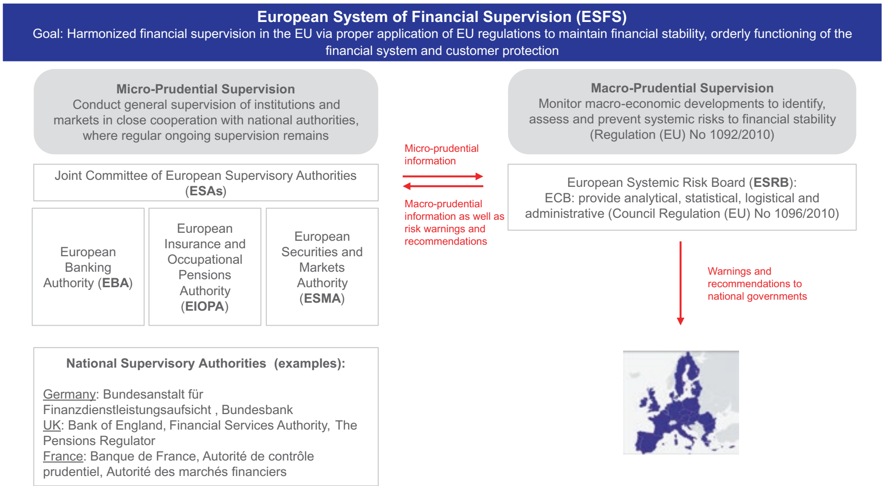


Fig. 9.5 European System of Financial Supervision ([16], p. 57)

Macro-prudential supervision[33] is provided by the ESRB, responsible for the monitoring of the macroeconomic environment to identify, assess, and prevent systemic risks[34] to financial stability. [35] In doing so, it cooperates closely with the ESAs as well as with respective international organizations, namely the IMF and FSB.

Micro-prudential supervision[36] is the specific supervision of financial institutions and individual markets by the ESAs in close cooperation with national authorities, each in charge of regular, ongoing supervisory practices. Three supranational organizations oversee micro-prudential supervision within the EU[37]:

- **European Banking Authority** (EBA)[38]: Guarantees an effective and consistent regulation and supervision of the European banking sector as well as the orderly functioning of the financial market.

[33] https://www.esrb.europa.eu/home/html/index.en.html, De Larosière, 2009: 46. Larosière (2009): 46.

[34] Systemic risk as defined by European law is the "*risk of disruption in the financial system with the potential to have serious negative consequences for the internal market and the real economy*" (Article 2(c) of Regulation (EU) No 1092/2010).

[35] ESRB Regulation : https://www.esrb.europa.eu/shared/pdf/ESRB-en.pdf?5020438634e48ca0076187a2a62e9344.

[36] De Larosière (2009): 48ff.

[37] http://www.eba.europa.eu/, https://eiopa.europa.eu/, http://www.esma.europa.eu/de.

[38] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010R1093&from=EN REGULATION (EU) No 1093/2010.

- **European Insurance and Occupational Pensions Authority** (EIOPA)[39]: Overseeing system stability and participant protection within the insurance and pensions market. In addition, it oversees the identification of market trends in order to address potential risk factors across sections and borders on a timely basis.
- **European Securities and Markets Authority** (ESMA)[40]: Responsible for securing the stability of the EU financial system. Fosters the integrity of the European securities markets as well as investor protection and convergence of supervisory structures between market regulators and the financial sector.

The ESAs, in close cooperation within a Joint Forum, monitor the transfer of European and international regulations into binding national laws in recommendations to national competent authorities of all EU member states. According to their statutes, their main objectives under consideration of consumer protection and common welfare are:

- Securing the safety, stability, and efficiency of the financial system
- Strengthening the **European Single Market**
- Providing for integral, transparent, and functioning financial markets

ESAs, with similarities to international standard setting bodies (SSBs), develop the technical standards required for a consistent implementation of European rules and regulations. The goal is to establish a *common rulebook* for all European member states to prevent regulatory arbitrage and fragmentation. These standards, categorized as *implementing* technical standards and *regulatory* technical standards, are subject to adoption (as decision or regulation) by the European Commission, Parliament, and Council. The former are amendments to "*nonessential elements of the legislative act*"; and the latter are specification "*where uniform conditions for implementing legally binding Union acts are needed*."[41]

The most important supervisory authority for exchange organizations within the EU is ESMA. ESMA's main responsibility is to guarantee investor protection and the stability and orderly functioning of financial markets. This is achieved by developing a single rulebook for financial markets and its consistent application throughout the EU. The latter is achieved either by guaranteeing supervisory convergence or by direct supervision of credit agencies and trade repositories. EMA's Board of Supervisors (which consists of representatives of all 28 member states, as well as third country observers and representatives of other European organizations) communicates all working products and developed regulatory decisions to the European Commission. Those decisions are usually provided by standing committees in support of several working groups of economic advisors, end consumers, and market participants.

[39] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010R1094&from=EN REGULATION (EU) No 1094/2010.

[40] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010R1095&from=EN REGULATION (EU) No 1095/2010.

[41] See [19], 395, based on Art 290f. TFEU.

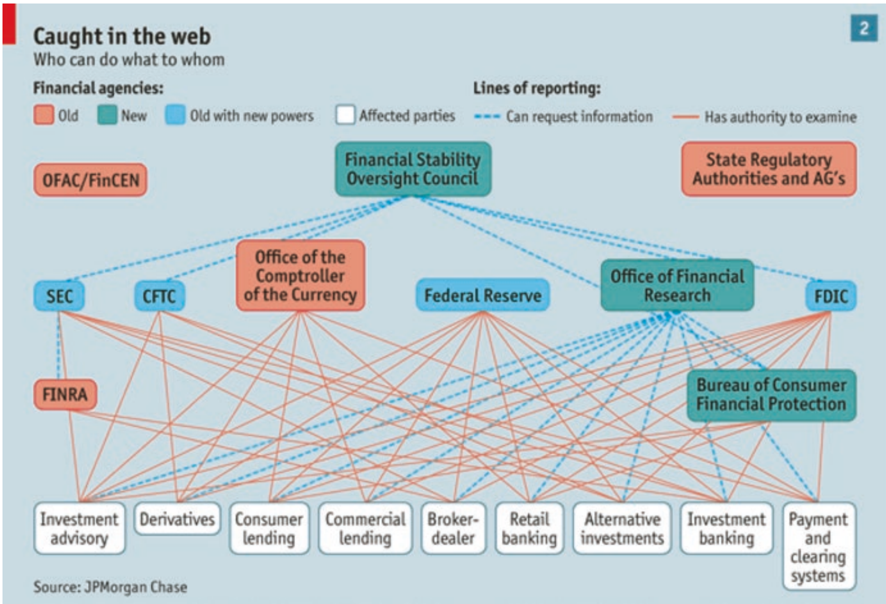


Fig. 9.6 Fragmentation and complexity of the US regulatory structure (source: JPMorgan Chase via The Dodd-Frank Act: Too big to fail. The Economist, February 15, 2012: http://www.economist.com/node/21547784)

9.2.1.5 US System of Financial Supervision

The US system of financial supervision remains fragmented (Fig. 9.6). Notwithstanding the level of integration of the US capital markets in comparison with the EU capital markets, that is evident across the financial industry sectors of banking, securities and futures, and insurance, as well as at the jurisdictional level between federal (national) regulators, and those within each of the 50 states for each subsector. Indeed, many financial sector actors are supervised by two or more authorities. The most recent proposal to significantly streamline the financial supervisory framework was started by the US Department of the Treasury, just before the onset of the Global Financial Crisis, and it was finally presented in March 2008.[42] Some of the proposed ideas were later part of the Dodd-Frank Act, adopted by the next administration in an election year turnover.

Most important, these included greater coordination among regulators through the Financial Stability Oversight Council (FSOC), and the extension of regulatory oversight to additional entities. These included systemically significant payment and settlement institutions and mortgage originators. The medium and longer term

[42] The Department of the Treasury Blueprint for a Modernized Financial Regulatory Structure (March 2008), available at http://www.treasury.gov/press-center/press-releases/Documents/Blueprint.pdf.

recommendations, however, have not been implemented to consolidate the overall number of US supervisors under a strengthened federal approach.[43]

The US financial regulatory structure remains significantly fragmented, particularly in the area of traditional deposit-taking activities of banks. There are multiple licensing and prudential supervisory authorities at both the federal level, and in each of the 50 states.[44] Banks may receive either a national license or a state license, and they are subject under each type of licence both the supervision of that state licensing authority, as well as to one of the two federal supervisory authorities, either the central bank or the deposit insurer.

Notwithstanding this, the substantive prudential rules applied have increasingly been harmonized, both in legislation and in their practical implementation. That has occurred through coordinating bodies that develop common supervisory and examination policies. In particular, the Federal Financial Institutions Examinations Council (FFIEC), established in 1979, promotes uniformity in the supervision of banking institutions.[45] Banking regulation in the USA has traditionally focused on prudence, with the applicable regulator overseeing banking risks to ensure that there are no disruptions to the credit cycle.[46]

Securities regulation, on the other hand, is typically focused on disclosure and mitigating conflicts of interest, fraud, and market manipulation. Banking regulation focuses on and oversees the capital and capital requirements of institutions. Securities regulation is written to ensure that market participants are provided with sufficient information to assess risks and make informed decisions.[47]

There are state regulators in the securities sector, but their responsibilities are more limited to smaller, more localized activities. The majority of securities markets activity falls under the Federal securities laws and the jurisdiction of the Federal Securities and Exchange Commission (SEC). Most notably, the futures industry in the USA, which had its origins in the agricultural sector, retains a distinct federal regulator, the Commodity Futures Trading Commission (CFTC). The powers of the CFTC were significantly enhanced through the Dodd-Frank Act.

---

[43] See id. at 14 (proposing consolidation by objective, resulting in three regulators responsible for market stability, prudential regulation, and business conduct, respectively).

[44] The majority of the 50 states themselves maintain regulators divided by the various sectors, although a minority have more recently introduced consolidated regulators within a state's jurisdiction. One prominent example is the Ney York State Department of Financial Services, which was created in 2011 by combining the state's previously independent banking and insurance supervisors. See www.dfs.ny.gov.

[45] The FFIEC recommends uniform principles, standards, and report forms for the federal examinations of financial institutions by the Federal Reserve Board, Federal Deposit Insurance Corporation, National Credit Union Administration and Office of the Comptroller of the Currency, and in relevant areas by the Consumer Financial Protection Bureau. See https://www.ffiec.gov/about.htm. See also the Conference of State Bank Supervisors, www.csbs.org.

[46] Congressional Research Service, Who Regulates Who and How? An Overview of the U.S. Financial Regulatory Policy, January 2015, Summary.

[47] Ibid. Summary.

The SEC has a clear and proven track record of enforcement actions to implement its mission of investor protection in addition to broader market oversight. The CFTC's traditional role has been much more like the EU's concept of investor protection, including in the MiFID legislation discussed below. In other words, the CFTC is focused on promoting the proper functioning of the markets.[48]

Regulation of the insurance industry is a matter of state competence, with a limited federal coordinating office established under the Dodd-Frank Act. Mortgage originators and money transmitters (as well as other financial service providers) are primarily regulated at the state level. The Federal Housing Finance Agency (FHFA) provides supervision over the US secondary mortgage market. The FHFA works to ensure that the housing government sponsored enterprises of Fannie Mae, Freddie Mac, and the Federal Home Loan Bank System provide liquidity and investment in the housing market in a safe and efficient manner.[49] The Consumer Financial Protection Bureau (CFPB) that was created by the Dodd-Frank Act, seeks to protect the American consumer, supervising certain retail banking and consumer products and services. It also enforces federal consumer financial laws.[50] Federal regulation for anti-money laundering and counter-terrorist financing is applied by the US Treasury Department's Financial Crimes Enforcement Network (FinCEN) for a broad range of federal and state-licensed financial services providers. All of the above-mentioned categories of regulated entities are included.[51]

The following are the major federal authorities in the US financial regulatory system (Fig. 9.7):

Federal Reserve System: Established in 1913 to provide stability to banks and trusts, the Federal Reserve System has three components: the Board of Governors of the Federal Reserve System, the Open Market Committee, and regional Federal Reserve Banks.[52] The Fed's Division of Banking Supervision and Regulation is responsible for the oversight of US bank holding companies, foreign banking organizations operating in the USA, state-chartered member banks of the Federal Reserve System, and any firm designated by the FSOC as systemically significant.

The Federal Reserve Board (FRB), which as of this writing is chaired by Janet Yellen, oversees the Federal Reserve Banks while setting national monetary policy and supervising the US banking system.

The Federal Open Market Committee (FOMC) is the Federal Reserve's primary policy-making body. The FOMC specifies the short-term objectives of open market operations, influencing the total money supply and credit available in the economy.[53]

---

[48] See James H. Freis Jr., An Outsider's Look into the Regulation of Insider Trading in Germany: A Guide toSecurities, Banking, and Market Reform in Finanzplatz Deutschland, 19 B.C. Int'l & Comp. L. Rev. 1, 77–79 (1996), http://lawdigitalcommons.bc.edu/iclr/vol19/iss1/2.

[49] See www.fhfa.gov.

[50] http://www.consumerfinance.gov/the-bureau/.

[51] See www.fincen.gov.

[52] Ibid. 23.

[53] http://www.federalreserve.gov/pubs/frseries/frseri2.htm.

Fig. 9.7 Overview of the Financial Stability Oversight Council

The 12 *Federal Reserve Banks*, located in Boston, New York, Philadelphia, Cleveland, Richmond, Atlanta, Chicago, St. Louis, Minneapolis, Kansas City, Dallas, and San Francisco are each responsible for the member banks located in their districts. These banks act as intermediaries between the member banks and the US federal banking system, both in supporting the payments system and in conducting supervisory functions delegated from the FRB.

Department of the Treasury: The Treasury Department is the executive department of the government responsible for promoting economic prosperity and ensuring the financial security of the U.S. The functions of the Treasury include, but are not limited to, managing federal finances; collecting taxes, duties and monies paid to and due to the USA and paying all bills of the USA; managing government accounts and the public debt; supervising national banks and thrift institutions; and advising on domestic and international financial, monetary, economic, trade, and tax policy.[54]

Office of the Comptroller of the Currency (OCC): The OCC is an independent bureau of the Department of the Treasury that licenses, regulates, and supervises national banks and federally chartered thrift institutions. Most major US banks are currently under an OCC national charter rather than a state charter.[55]

Federal Deposit Insurance Corporation (FDIC): The FDIC regulates federally insured depository institutions, including state banks and thrifts that are not members of the Federal Reserve System. The FDIC was created to offer small depositors the comfort of a "guarantee," so that if their bank failed, the depositor would not lose the full amount of their deposits.[56] Dodd-Frank expanded the FDIC's role in the liquidation of non-bank entities with systemic significance.

The National Credit Union Administration (NCUA): The NCUA oversees federally chartered or insured credit unions. [57] The NCUA, with the full backing of the US Government, operates and manages the National Credit Union Share Insurance Fund (NCUSIF), an entity analogous to the deposit insurance for banks overseen by the FDIC.[58]

Securities and Exchange Commission (SEC): The enactment of the Securities Exchange Act of 1934 created the SEC as a government agency to restore confidence in the US markets after the Great Depression. As detailed in its mission statement, the SEC was created "*to protect investors, maintain fair, orderly, and efficient markets, and facilitate capital formation.*"[59] The SEC oversees the US securities markets, regulating a wide array of entities including securities exchanges, brokers and dealers, clearing agencies, mutual funds, investment advisors (including entities commonly known as hedge funds with over $150 million in assets), security-based swap dealers (SBSD), major security-based swap participants (MSBSP), and security-based swap execution facilities (SB SEF). The SEC also oversees corporations that are selling securities to the public.

Commodity Futures Trading Commission (CFTC): The CFTC was formed in 1974 to provide a regulatory framework for the derivatives market. The derivatives market has increasingly been becoming more complex. The CFTC regulates futures exchanges, brokers, commodity pool operators, commodity trading advisor, as well as swap dealers (SD), major swap participants (MSP), and swap execution facilities (SEF). The objective of the CFTC is to "*foster open, transparent, competitive, and financially sound markets, to avoid systemic risk, and to protect the market users and their funds, consumers, and the public from fraud, manipulation, and abusive practices related to derivatives and other products that are subject to the Commodity Exchange Act.*"[60]

---

[54] http://www.treasury.gov/about/role-of-treasury/Pages/default.aspx.

[55] http://www.occ.gov/about/what-we-do/mission/index-about.html.

[56] Congressional Research Service, Who Regulates Who and How? An Overview of the U.S. Financial Regulatory Policy, January 2015, 21–22.

[57] Ibid, 23.

[58] http://www.ncua.gov/about/Pages/default.aspx.

[59] http://www.sec.gov/about/whatwedo.shtml.

[60] .http://www.cftc.gov/About/MissionResponsibilities/index.htm.

Financial Stability Oversight Council (FSOC): The FSOC (which we describe in more detail later in the chapter) is tasked by the Dodd-Frank Act with monitoring the financial stability of the US financial system by identifying and responding to risks and by promoting market discipline.[61] The FSOC has the ability to designate financial institutions as being "systematically important." When so designated, these entities are subject to a greater degree of scrutiny and oversight. The Financial Stability Oversight Council has ten voting members and five non-voting members:

The nonvoting members are representatives from the Office of Financial Research (OFR) as well as the Federal Insurance Office, a state insurance commissioner, a state bank supervisor, and a state securities commissioner.

9.2.2 Global Regulatory Processes

The term financial market regulation describes the legislative processes underlying the previously outlined supervisory processes. In the wake of the systemic weaknesses revealed through the financial crisis a reformation of the regulatory environment of financial markets was started in 2008. The establishment of a new order within the system (*new normal*) was the objective. The shortcomings included gaps in regulation and supervision, ineffective risk management, the lack of transparency, and problematic behavior, products, services, and strategies of some market participants.

In summary, the reform of financial market regulation focuses on four integral points:

- Safety: Specific definition of responsibilities especially regarding risk management
- Integrity: Prevention of unnecessary risk potentials for organizational and systemic integrity
- Efficiency and transparency: achieved by simplified market and product structures as well as by extensive reporting and trading obligations, and a consistent burden-sharing and protection of taxpayer's money

The G20, which is based on these objectives, developed its first agenda for the reform of global financial markets in 2009. Its key points include a higher capital basis for financial institutions; reforming the OTC market and enhancing the supervision of systemically important financial organizations. The revised 2014 agenda focuses on strengthening the overall stability of financial institutions, the ending of **too-big-to-fail**[62] policy, managing the shadow banking system, and, finally,

---

[61] http://www.treasury.gov/initiatives/fsoc/Pages/home.aspx.

[62] See FT Lexicon: This term is most often used to denote banks and firms that would substantially damage the financial system and the rest of the economy should they "fail," i.e., go bankrupt.

The logic is therefore that these organizations would receive a "bailout" of some kind from the government—at the very least, protecting creditors against losses and perhaps also enabling management to stay in place (and, in some cases, the full payment of wages and bonuses).

increasing the stability of the derivatives market. The overriding principles for targeted the new order of the financial system continue to be investor and system protection.

G20 leaders, alongside the review of the general framework of financial market regulation and supervision, emphasized the importance of international cooperation in pursuit of a level playing field in the development and implementation of policies in the legislative process.

9.2.2.1 International Level

A regulatory initiative passes through various stages until its final implementation as national law. The cascade is exemplarily described using OTC Derivatives Clearing, an already implemented G20 2009 agenda point (cf. Fig. 9.8):

The FSB, as the main surveillance organ overseeing the implementation of the G20 agenda, publishes so-called *Progress Reports* on a regular basis. The reports summarize the status of implementation of the G20 agenda. The SSBs (1) BCBS and (2) IOSCO, in standards and recommendations, firm up requirements that are set for supranational and international regulation. Those reports are a basis for national/European legislation, and they provide a guideline for the specific rulemaking process. Indeed, they are of particular importance for the development of the proverbial level playing field in the implementation of international rules and regulations.[63]

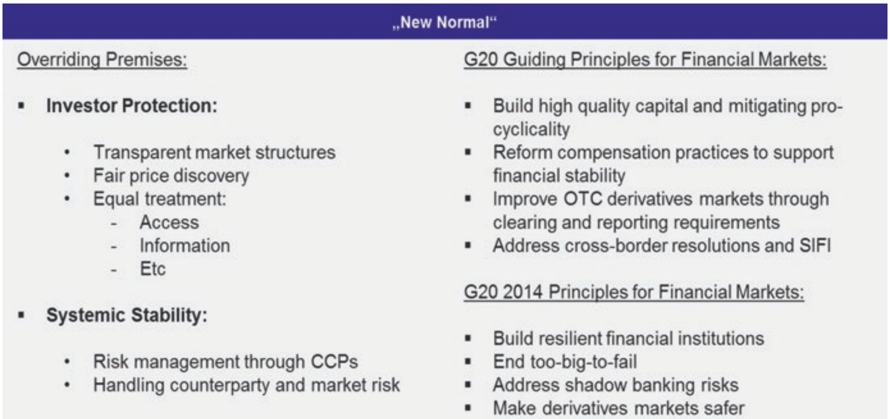


Fig. 9.8 "New Normal": Premises and Principles (http://www.bundesregierung.de/Content/DE/StatischeSeiten/Breg/G8G20/Anlagen/G20-erklaerung-pittsburgh-2009-en.pdf?__blob=publicationFile&v=2; See [12] http://www.g20australia.org/g20_priorities)

[63] See [13], p. 475.

A main problem facing international regulation is **regulatory arbitrage**. This is especially so in cases that (1) allow for a certain content and/or scope of amendment during the national implementation process (if, say, only minimum standards are defined by legislation, e.g., when implementing regulatory standards through a directive rather than a regulation on the European level); or (2) where the implementation process has significantly different time scopes. For example, in the EU with rather slow proceedings compared to the USA where a majority of the required regulations being included in Dodd-Frank are already enacted.

Both problems are connected to the example laid out in Fig. 9.8: Dodd-Frank covers some aspects of trading and clearing, whereas EMIR covers clearing and reporting only. Final implementation of EMIR, due to the prevailing variety of national regulatory systems, takes somewhat longer in the EU than it does in the USA. While EMIR was still in the legislative process, Dodd-Frank was already enacted. The delayed introduction of EMIR has allowed US financial firms to offer the resulting services faster in Europe than their European competitors. For instance Derivatives-OTC-Clearing was implemented in the USA in 2013; by contrast, implementation in the EU follows a phase-in approach as from 21 June 2016.[64]

9.2.2.2 European Rulemaking Process

Figure 9.9 provides a general overview of the common European rulemaking process. On financial markets regulation specifically, a simplified legislative procedure—the **Lamfalussy process**—was introduced in 2002 to accelerate the complex process by the application of a four stages plan. On *level 1* EU institutions develop the framework legislation initiated by the Commission. The Commission, supported by expert committees comprised of representatives of the member state's finance ministries,[65] and the European Supervisory Authorities (ESAs), then prepares technical standards for the application of the legislation (*level two*). Technical standards are distinguished by Regulatory Technical Standards (RTS), adopted by the Commission through a Delegated Act—meaning that "*the legislator delegates the power to adopt acts amending non essential elements of a legislative act to the Commission*"—and Implementing Technical Standards (ITS), adopted by means of an Implementing Act—meaning that "*European measures which require uniform implementation in the Member States directly authorise the Commission to adopt*" those acts.[66]

The implementation measures are defined by the Commission during *level 3*, once again under the support of specialized committees with members of the

[64] http://www.efinancialnews.com/story/2015-05-29/europe-to-begin-derivatives-clearing-from-spring-2016.

[65] Namely: European Banking Committee (EBC), European Securities Committee (ESC), European Insurance and Occupational Pensions Committee (EIOPC) and Financial Conglomerates Committee (FCC).

[66] European Union legal acts: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=URISERV:ai0032&from=EN.

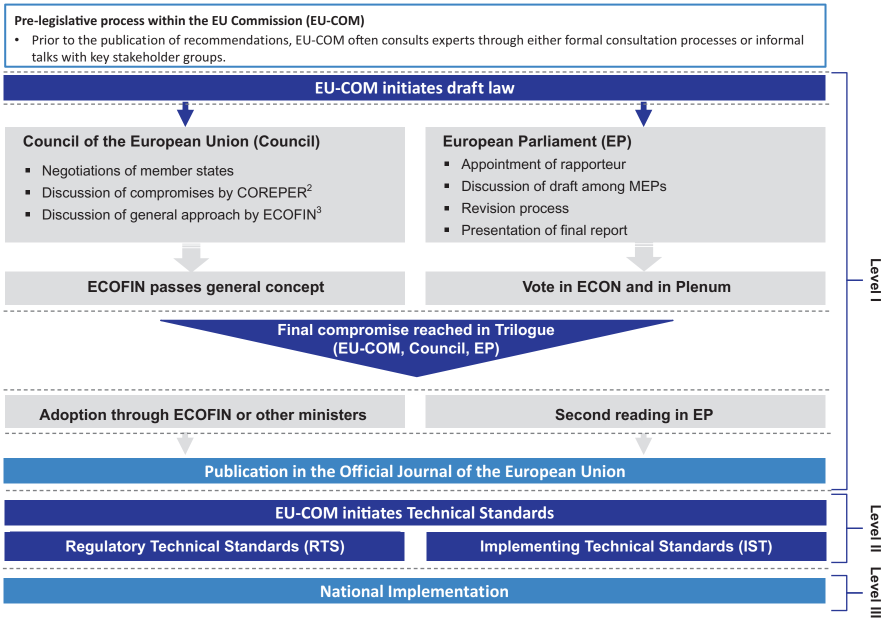


Fig. 9.9 The European rulemaking process

national supervisory authorities and central banks.[67] The objective is to achieve harmonized and consistent legislation and supervision. In the post-implementation phase, the appropriate application of Community Law in the member states is verified and guaranteed.

In contrast to other European legislations that are defined by *minimal harmonization* rather than *mutual recognition*,[68] most of the Directives and Regulations covering financial markets follow the concept of *maximum harmonization*[69] in light of

[67] Committee of European Banking Supervisors (CEBS), Committee of European Securities Regulators (CESR), Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS).

[68] Regulation (EC) No 764/2008: *"The principle of mutual recognition, which derives from the case-law of the Court of Justice of the European Communities, is one of the means of ensuring the free movement of goods within the internal market. Mutual recognition applies to products which are not subject to Community harmonisation legislation, or to aspects of products falling outside the scope of such legislation. According to that principle, a Member State may not prohibit the sale on its territory of products which are lawfully marketed in another Member State, even where those products were manufactured in accordance with technical rules different from those to which domestic products are subject."*

[69] The principle of full harmonization was first laid out by the Court of Justice of the European Communities as well. In its judgment regarding Joined Cases C-261/07 and C-299/07 on Directive 2005/29/EC (Unfair Commercial Practice) it is concluded that: *"The Directive is based on the principle of full harmonisation. This means that Member States can no longer implement or apply either less or more restrictive or prescriptive consumer protection measures in the area it harmonises. As the Preamble to the Directive explains, in order to remove internal market barriers*

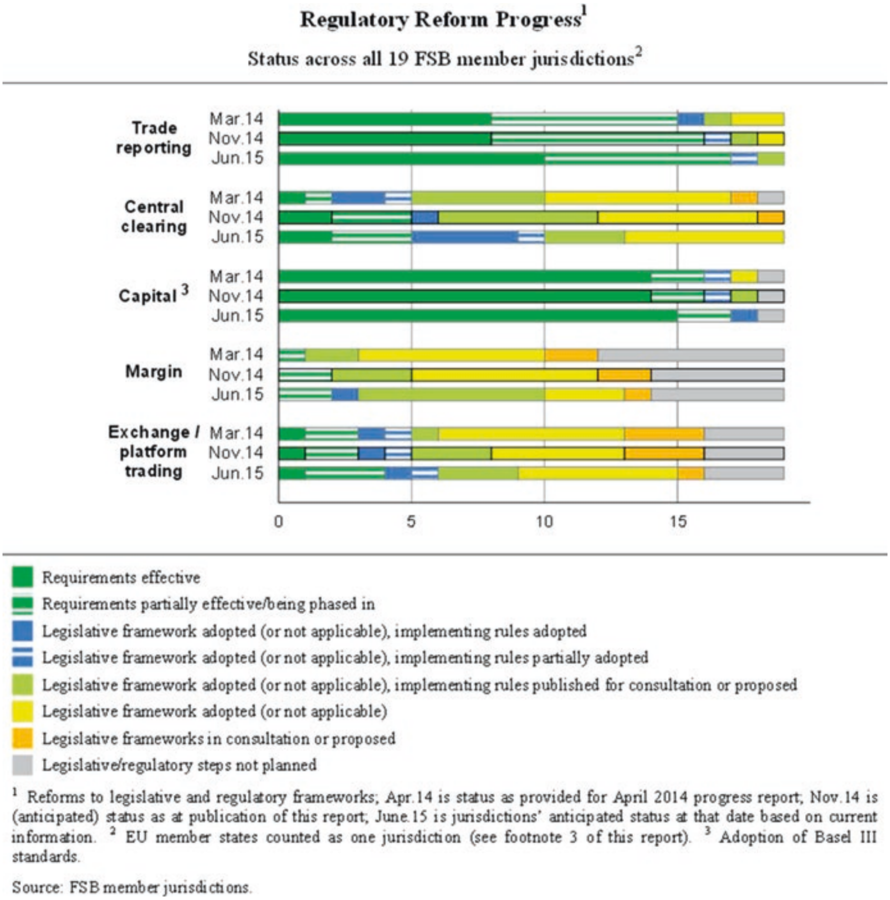


Fig. 9.10 Global rulemaking cascade—OTC Derivatives Clearing

enforcing "a level playing field" for implementation and regulatory practices within member states.

Another characteristic common to financial market regulations in the EU is the objective of financial stability. That is achieved through safe and sound financial markets characterized by transparent practices, institutions and instruments, and a high level of consumer protection.

The legislative process often takes years from pre-drafting until implementation (Fig. 9.10). For example, the **Markets in Financial Instruments Directive (MiFID)** took more than 4 years from pre-drafting to its level 1 agreement (2000–

---

caused by regulatory disparities and to increase legal certainty for both consumers and businesses, it was necessary to replace existing national systems with a uniform regulatory framework at Community level." (https://webgate.ec.europa.eu/ucp/public/index.cfm?event=public.guidance.showArticle&elemID=52).

2004). Another example is EMIR, where the process took 3 years (July 2009–July 2012). The technical implementation in level 2 is still continuing in 2016.

For the upcoming years, in particular when most of the initial G20 agenda is implemented, the industry will likely face reviews and the fixing of existing regulation. This was certainly the case with the MiFID review which was finalized and published in 2014. The work on global regulatory convergence and coherence, as well as **third country rules**, is expected to be a focal point in the years to come.

9.2.2.3 US Rulemaking Process

In comparison to the EU, the process in the USA for the development of financial regulation is relatively straightforward, consisting of the law-making process followed by implementation through rulemaking.

Legislation in the USA generally may be initiated in either of the two US Houses of Congress—the House of Representatives or the Senate. A proposal from the administration, which on financial matters would normally be developed by the Department of the Treasury, is coordinated by the White House. A Member of Congress usually from the same political party as the administration then requests a legislative bill be introduced.

Bills are considered by the relevant competent committee, such as the Senate Banking Committee or the House Financial Services Committee. If the bill is passed by them, it is then referred to the full House. The bill, to become law, must be passed by each House of Congress—often worked out in a process to reconcile differences—and then finally signed into law by the President.

Most legislative bills that are introduced do not become law, and legislation that eventually does pass into law often involves multiple iterations by both Houses of Congress, the Administration, and contributions from the general public. A notable feature of the US legislative process includes the distinction between the processes and the responsible committees within the Congress for the authorization of activity, and the appropriation of funding to carry it out.

Hence, it is possible to have activity mandated, but not funded for implementation. Moreover, funding may be restricted or withheld with certain activities. US legislation, particularly in the area of financial regulation, generally establishes principles and establishes authority or competence for the administration, or an independent regulatory authority to implement through detail rulemaking. These include the Board of Governors of the Federal Reserve System, SEC, or CFTC. With limited exceptions, rulemaking is developed in accordance with the Administrative Procedure Act.[70] Most notably, this requires that rules be proposed for public comment. The comments must be reviewed and considered before finalization and implementation of the rule. Even financial regulations are increasingly subject to a number of cost-benefit analyses, both within the administration and as part of the formal public justification for the rulemaking.

[70] U. U.S.C. § 500 et seq.

9.3 Effects of Regulation on the Value Chain of Market Operators

Regulation has had far-reaching global significance since the financial crisis in 2007 and 2008, and the G20 principles that followed. The regulatory institutions and initiatives, both across borders and markets, today shape the business and strategic approach of exchange organizations. Chapter 2 depicts a non-exhaustive list of these initiatives, and it schematically outlines how, on many different levels, an integrated market operator in the EU can be affected by regulation. We here focus on the value chain of Deutsche Börse Group as an illustration.

EU and US regulators established a Financial Market Regulatory Dialogue in 2002 as a platform for deliberating regulatory developments and for exchanging information on new legislative initiatives in order to "*implement and enforce robust standards, including those on the G20 financial regulatory agenda.*"[71]

The following is a high-level overview of the most important regulations for exchange organizations that have the most lasting impact today. The details of named policies should be gleaned from the official legislative texts and amending documents. The recommendations for detailed assessments are provided in footnotes.

9.3.1 Markets in Financial Instruments Directive (MiFID I; MiFID II/MiFIR)[72]

MiFID is one of the most comprehensive legislations in the framework of European financial markets regulation. This legislation, developed with superlative precision to harmonize financial markets within the European Single Market, promotes the integration, competitiveness, and efficiency of EU financial markets via the Directive.

MiFID I, adopted in April 2004, replaced the Investment Services Directive (ISD) operating since 1993. MiFID I was developed as a regulatory response to significant changes in the overall financial market structure. The changes were triggered by continuous electronization of trading, as highly complex and diversified financial instruments and trading services were introduced, and disintermediation followed. EU-COM agreed in 2006 to the implementation measures. Member states had until 2007 to implement them.

---

[71] United States-European Union Financial Markets Regulatory Dialogue Joint Statement of Jan 2015: http://ec.europa.eu/finance/general-policy/docs/global/150115-us-eu-joint-statement_en.pdf.

[72] The Directive is available here [7]: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02004L0039-20110104&from=EN.

For a detailed analysis see http://safe-frankfurt.de/uploads/media/Gomber_Nassauer_MiFIDII_MiFIR.pdf.

MiFID regulates all trading venues and trading practices of financial instruments, especially equities. The major components encompass extensive developments of the EU financial market structure, including[73]:

- Chapter I: Conditions and Procedures for Authorization:
  - Art 5 in conjunction with Art 4—definition, classification, and authorization of trading venues: Regulated Markets vs. Multilateral Trading Facilities (MTF) vs. Systematic Internalizers (SI).[74]
  - Art 15—nondiscriminatory access to financial markets induced by transparency, equal treatment, and neutrality, especially facilitating cross-border financial services.
- Chapter II: Operating Conditions for Investment Firms:
  - Regarding investor protection: Art 19—enhanced investor protection defined by client focus, transparency, fair and equal treatment in conjunction with Art 21—definition of Best Execution practices.
  - Regarding market transparency and integrity: Art 25—transaction recording and reporting and Art 28f. pre- and post-trade transparency requirements especially regarding MTFs.
- Chapter III: Rights of Investment Firms:
  - Art 31—increased competitiveness in investment services, Art 32—market harmonization.
- Chapter IV: Regulated Markets:
  - Art 33f.—nondiscriminatory access, Art 44ff.—pre-trade, post-trade transparency and post-trade infrastructure regulation.

Reporting requirements extended, while exchanges already provided high level of pre- and post-trade transparency an enhanced transparency regime for equities was introduced to facilitate implementation of Best Execution.[75]

At its core, MiFID has had a considerable impact on the infrastructure of equity markets, significantly changing market mechanisms, products, and offerings.

As the academic literature clearly points out, competition for order flow between trading venues has improved, and implicit costs (spreads and market impact) as well as explicit trading costs (fees) have decreased considerably.[76] However, with this

---

[73] Directive2004/39/EC:http://zeur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02004L0039-20110104&from=DE.

[74] Systematic internalizer as defined by the Commission *"means an investment firm which, on an organised, frequent and systematic basis, deals on own account by executing client orders outside a regulated market or an MTF"* (EU Directive 2004/39 EC (Art 4, Abs 1, Satz 7)).

[75] WpHG online: http://www.gesetze-im-internet.de/wphg/__33a.html.

[76] Peter Gomber/Axel Pierron, 2010: MiFID—Spirit and Reality of a European Financial Markets Directive; Peter Gomber/Benedikt Jäger, 2014 [8]: MiFID: Eine systematische Analyse der Zielerreichung.

focus on market integrity and investor protection, the outcome of MiFID I show it has not completely fulfilled its objectives.

One of the major shortfalls of MiFID I was the introduction of a clear delineation between regulated trading venues and OTC trading. The most important consequence is a comparably large share of OTC trading and a lack of data quality. The mitigation of market impact that results from large orders—often stated as a central motivation for OTC trading—is not deducible from available data. Large orders are commonly broken up into smaller retail size pieces that can be executed more easily without price impact in a transparent and liquid market.

In 2010, with the lessons learned from the crisis, hearings and consultations on MiFID started, with a revision of the established Directive previously anticipated in Art 65 MiFID. The review aimed at further increasing market efficiency, transparency, integrity, and resilience with a focus on [77]

- Expanding trading obligations for financial instruments and introducing **organized trading facilities (OTF)**. In addition, the requirement that certain OTC derivatives be traded on multilateral trading venues was introduced. This secured that, apart from defined exceptions, trading takes place on a regulated trading facility.
- Increasing existing transparency requirements and adjusting them to cover non-equity instruments and markets.
- Improving market access for third-country firms based on equivalent assessments.

In October 2011, EU-COM adopted a proposal for a Directive[78] amending MiFID I as well as a complementing Regulation (MiFIR[79]). The respective agreement of EU Parliament and Council was made in January 2014. The legislative framework was published in the EU Official Journal in June 2014. Member states had a 2-year period to transfer the rules into state law, to take effect in January 2017.[80]

In the USA, one notable regulation in the structure of exchanges and the development of electronic forms of trading is the SEC's rulemaking on the National Market System[81] introduced in 2005, and generally referred to as Regulation NMS (Reg NMS). The Regulation implements Section 11A of the Securities Exchange Act requiring the SEC to promote five objectives through the National Market System: (1) efficient execution of securities transactions; (2) fair competition among broker-dealers, among exchange markets, and between exchange markets and non-

[77] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011SC1227&from=DE, pg. 5ff. SEC (2011) 1227 final: 5ff.

[78] Directive 2014/65/EU: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2014:173:FULL&from=EN, pg. 351ff. Directive 2014/65/EU: 351ff.

[79] Regulation (EU) 600/2014: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2014:173:FULL&from=EN, pg. 85ff.

[80] http://ec.europa.eu/finance/securities/isd/mifid2/index_de.htm, EC Press Release of 12 June 2014: http://europa.eu/rapid/midday-express-12-06-2014.htm?locale=en

[81] 70 FR 37,496 (29 June 2005).

exchange markets; (3) price transparency; (4) best execution of investor orders; and (5) an opportunity, consistent with the foregoing, for investor orders to meet without the participation of a dealer.[82] The new substantive rules were designed to modernize and strengthen the regulatory structure of the US equity markets.[83] Rule 611, known as the Order Protection Rule, generally requires that trading centers establish, maintain, and enforce written policies and procedures reasonably designed to prevent the execution of trades at prices inferior to protected quotations displayed by other trading centers. Rule 610, known as the "Access Rule," requires fair and nondiscriminatory access to quotations, and establishes a limit on access fees to harmonize the pricing of quotations across different trading centers.[84] The overall goal of Reg NMS is to provide investors best price execution for their order by encouraging competition among executive venues. The regulation, however, has also been criticized for contributing to excessive fragmentation in the US markets.

9.3.2 European Market Infrastructure Regulation[85]

The European Market Infrastructure Regulation (EMIR), on the heels of decisions at the G20-Pittsburgh Summit, aims at regulating the OTC Derivatives Market and to reduce systemic risks by regulatory harmonization and by increasing the safety and efficiency of the European post-trade landscape. EMIR facilitates transparency, the reduction of counterparty credit risk, and the decrease of operational risks associated with OTC derivative trading. This regulation also promotes the standardization of OTC derivatives contracts and their processing.

EMIR was adopted by the EU Parliament and Council in 2012 (level 1). The initial set of (level 2) standards, the technical details, was enacted on March 2013. The first European trade repositories were approved by ESMA in late 2013. Since 2014, the reporting of OTC trades, as well as the authorization of European clearing houses, has been mandatory, with the final clearing obligation taking effect as from 21 June 2016.

The Regulation applies to all central counterparties as well as to their members, financial counterparties[86] and **trade repositories**. The most important components are:

---

[82] See 15 U.S.C. § 78k-1.

[83] See70 FR. at 37,496.

[84] See SEC Division of Trading and Markets, Responses to Frequently Asked Questions Concerning Rule 611 and Rule 610 of Regulation NMS (4 April 2008 Update), available at https://www.sec.gov/divisions/marketreg/nmsfaq610-11.htm.

[85] Regulation available online: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32012R0648&from=DE, details on specific aspects of the Regulation and according application are provided in Chap. 5 (Clearing).

[86] According to Regulation (EU) No 648/2012 financial counterparties include investment firms, credit institutions, insurance/assurance/reinsurance undertakings, occupational retirement provisions, and alternative investment funds.

- Art 4—Clearing obligation:
  - Standardized OTC derivatives must be CCP cleared.
  - ESMA/European Commission are responsible for final decision on eligibility.
- Art 7, 8—Access requirements from trading venues to CCPs and vice versa:
  - Trading venues must have nondiscriminatory access to CCP derivatives clearing services for OTC derivatives, unbundled from trading, for eligible OTC derivatives and vice versa from CCPs to trading venues.
- Art 9—Reporting obligation:
  - All market participants are required to deliver details regarding concluded, modified or terminated derivative contracts to a trade repository.
- Art 14ff.—Authorization and Supervision of CCPs:
  - Authorization is processed on the national level with the CCP's competent authority of the respective member state and in course of a session of the College of Supervisors, including ESMA and other national competent authorities.
- Art 25ff.—recognition of third-country CCPs in terms of allowing non-EU CCPs to offer services in EU with the requirement to be EMIR compliant. Most notable in this regard was the common approach agreed in February 2016 between the European Commission and the US CFTC after multiple years of negotiations.[87]
- Art 26ff.—CCP requirements:
  - Regulation introduces requirements on conduct of business, organizational rules, and risk management including, among others, split of pricing for trading and clearing, installment of risk committees dominated by clearing members, and default funds to cover the possible failure of two of the largest clearing members.
- Art 39—Account segregation (offering at least omnibus client segregation and individual client segregation).
- Art 51ff.—Interoperability:
  - Interoperability to be granted among CCPs for money market instruments and transferable securities.

As required by Article 85 (1) of EMIR, the EU Commission started to undertake a review of the Regulation in April 2015. The review will pay special attention to the transfer of systemic risk to CCPs as provided for by the general clearing obligation included in EMIR, increasing the rate of centrally cleared OTC derivative transactions. The risk of failure of the CCP itself has to be accounted for. That is because

[87] See CFTC Press Release of 10 February 2016, http://www.cftc.gov/PressRoom/PressReleases/cftc_euapproach021016.

this shift not only improved risk management mechanisms for clearing members, but also increased the concentration of risk within the respective CCP as well as structural interdependencies between the CCP and its members and interoperating CCPs. Although the probability of such failure is relatively low, EU regulators still demand thorough recovery and resolution provisions for CCPs to manage systemic risk above and beyond the general risk management mechanisms required by EMIR. Hence, it can be expected that the EMIR review and recovery and resolution requirements will once again change the organization of CCPs.[88]

9.3.3 Central Securities Depository Regulation[89]

On March 7, 2012, the European Commission released its legislative proposal to the Council and Parliament for the reform of the European settlement landscape and the creation of new harmonized rules for Central Securities Depositories (CSD) on processes, organization, and conduct. The Central Securities Depositories Regulation (CSD-R) promotes defragmentation of the settlement business, reduction of cross-border issues, a decrease of barriers to CSDs, and lower risks and costs for settlement activities.

On August 28, 2014, the final CSD-R was published in the Official Journal, and it then became effective the following month. The complete application, including the CSD authorisation and compliance requirement, is slated for end of 2016 or latest early 2017.

The CSD-R is applicable to all activities of CSDs in the settlement of financial instruments within the EU (Art 1). The proposed regulation covers settlement discipline (Art 6ff.) and sanctions for the overall market, with measures for the harmonization and improvement of CSD services in the EU (Art 33ff.). This single settlement regime should facilitate systemic stability and, at the same time, decrease costs and risks of settlement fails from fragmentation and an unlevel playing field that results from the diverse rules applicable to settlement processes within the EU. CSD-R aims to provide CSDs with a single *European license* that shares the same conditions for authorisation in all EU countries (Art 10ff.), and also the same supervisory processes (namely (a) on a national level through the competent authority and (b) on a community level through ESMA).

Moreover, the regime imposes measures to decrease barriers in the access to and interoperability of CSDs (Art 49) between CSDs (Art 50ff.) and between CSDs and FMIs (Art 53) to foster stability, efficiency, and competition for better investor protection. Finally, it imposes prudential requirements (Art 42ff.) largely inspired by CPSS-IOSCO as well as the harmonization of settlement in establishing a single settlement cycle for the EU. The latter requires that the execution of settlement take

[88] For details please refer to Chap. 5 Clearing.

[89] Regulation (EU) No 909/2014: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0909&from=EN.

place no later than two business days after the respective trade (so called $T+2$: Art 5), a measure of special importance for cross-border settlement.

CSD-R, due to its harmonization effect, significantly supports the European settlement landscape for the introduction of TARGET2-Securities (T2S), the single pan-European platform for securities settlement in central bank money. T2S is considered a catalyzer for the objectives of the CSD-R, namely harmonization, competition, system stability, and reduced costs, as well as increased customer protection. The sweeping nature of these changes and the unavoidable investment in infrastructure necessary to achieve them may, in the short-term, limit the available savings for individual settlement transactions. Banks, nevertheless, in the long term, are expected to realise significant capital, funding and operating savings by delayering[90] and consolidating security and cash holdings.[91]

CSD-R in Europe shares one significant goal in common with developments in the US, albeit through industry initiative rather than legal prescription: The move from a $T+3$ to $T+2$ settlement will help mitigate counterparty risk, reduce costs, and optimize allocation of capital.[92] The designation of DTC as a systemically significant financial market utility subjected this industry utility to regulation under the Dodd-Frank Act, as described below.

9.3.4 Capital Requirements Directive IV (CRD IV)[93]

The European implementation of the Global Regulatory Framework for Capital and Liquidity (Basel III) is covered by the CRD IV Package containing the Capital Rights Directive (CRD) and Regulation (CRR). The Package was adopted in summer 2013 and enacted on January 1, 2014, with full implementation required by January 1, 2019. CRR contains comprehensive prudential and liquidity requirements for credit institutions and investment firms (esp. Single Rule Book). CRD covers access to deposit-taking activities and requirements in national supervisory authorities. EBA is mandated with the consultation and issuance of Level II texts and with the development of Regulatory and Implementing Technical Standards.

The most important provisions of the initiative:

[90] Delayering stands for reducing intermediation or reducing the necessary layers needed to access settlement in central bank money cross border.

[91] Clearstream, the T2S Opportunity: Unlocking the Hidden Benefits of Target2 Securities, http://www.clearstream.com/blob/68228/9f9261051598b77e44bddf291d655859/t2opportunity-pdf-data.pdf.

[92] DTCC Recommends Shortening the US Trade Settlement Cycle (April 2014), http://www.dtcc.com/~/media/Files/Downloads/WhitePapers/t2-Shortened-Cycle-WP.pdf.

[93] Regulation (EU) No 575/2013: http://eurlex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32013R0575&from=EN.

Directive 2013/36/EU: http://eurlex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32013L0036&from=EN.

- Enhanced risk management among financial institutions
- Increased capitalization requirements for trading derivatives (especially OTC) and banks' exposure to CCPs in the form of default fund contributions
- Introduction of a leverage ratio and of liquidity standards based on Basel III
- Revised capital definition and capital ratios including provisions for systemically important banks
- Enhanced reporting requirements and tighter corporate governance/remuneration rules

9.3.5 *Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act)*

The dramatic events of the Global Financial Crisis from 2007 to 2008 had the most brutal impact on the US financial markets since the Great Depression of the 1930s. As with many previous financial crises, this Global Financial Crisis brought calls for changes in the regulatory system, and prompted financial reform legislation. In the end, a variety of measures coalesced into a legislative package signed into law as the Wall Street Reform and Consumer Protection Act of 2009, commonly known as the Dodd-Frank Act (the "Act")[94] after its two principal sponsors in the US Congress—Senate Banking Committee Chairman Chris Dodd and the House of Representatives Financial Services Committee Chairman Barney Frank.

President Barack Obama signed Dodd-Frank into law on July 10, 2010. The Act consists of 848 pages in its official published version, and it is categorized by 16 chapters known as "titles." With the overall objective of promoting the financial stability of the USA, Dodd-Frank seeks to improve accountability and transparency in the financial system, to end "too big to fail," and by ending bailouts to protect the American taxpayer and consumers from abusive financial services practices.

Among other things, the Dodd-Frank Act created a new entity, the Financial Stability Oversight Council (FSOC), to identify certain financial firms as "systematically important" and to then subject them to greater regulation.

The Act further consolidated bank supervision by merging the Office of Thrift Supervision (OTS) into the Office of the Comptroller of Currency (OCC). The Act requires more derivatives to be cleared and traded through exchanges, and it creates new categories of market participants who are subject to registration and reporting requirements.

Sweeping changes to bank capital, securities, derivatives, systemic risk, executive compensation, bank activities, and liquidation rules are the critical components of the Dodd-Frank Act that have led to the stabilization of the US capital markets. Below, we outline these titles in the Act:

[94] Pub. L. 111-203, 124 Stat. 1326-2223. As a type of omnibus legislation, the individual provisions of the law are codified across many different provisions of the US Code, including as amendments within different provisions of the preexisting banking and securities laws.

Title I—Financial Stability

- Title I established the Financial Stability Oversight Council (FSOC) and Office of Financial Research (OFR), to promote financial stability in the USA.
- FSOC, chaired by the Secretary of the Treasury, monitors the stability of the US financial system by identifying risks and responding to emerging risks and promoting market discipline.[95] The Council facilitates information sharing and coordination among member (see above) and other federal and state agencies as well as acting as an adviser to the financial regulatory agencies. FSOC recommends new or enhanced standards for activities that increase risks within the US financial markets. Under the terms of the Act, the FSOC shall consult with appropriate foreign regulatory authorities in exercising its duties with foreign financial companies, cross-border activities, and markets.[96]
- FSOC is mandated with identifying "systematically important" non-bank financial companies, a designation that puts these entities under the regulation of the Federal Reserve. In other words, under higher regulatory standards. Systematically important entities are identified as those that could pose significant threats to financial stability through company activities. In designating entities as systematically important, FSOC considers factors such as leverage, the assets and liabilities of the company, company activities, and assets under management in addition to any other factors that are considered applicable by the FSOC.
- Over the period July 2013 through December 2014, FSOC designated three non-bank financial companies as systemically significant: MetLife, General Electric Capital Corporation, and Prudential Financial. These were three entities that had primary roots in the insurance industry, with GE Capital as the financing arm for the well-known diversified manufacturing conglomerate.[97] MetLife has contested that designation, and GE subsequently spun off its financial operations.[98]
- The OFR, operating as a new office within the US Department of the Treasury, is the information gathering arm of FSOC. The office spans the financial system to measure and analyze risks, perform essential research, and to collect and standardize financial data.[99] FSOC, acting through the OFR, can request reports from any financial entity to evaluate potential threats to the US financial marketplace.
- Section 165 of Dodd-Frank requires that companies covered by this law create and submit a Resolution Plan to the FSOC, FDIC, FRB. The Plan should focus on the company's approach to an orderly resolution during bankruptcy. It should

[95] http://www.treasury.gov/initiatives/fsoc/Pages/home.aspx.

[96] Dodd-Frank Act § 113(i).

[97] http://www.treasury.gov/initiatives/fsoc/designations/Pages/default.aspx.

[98] See MetLife, Press Release, MetLife to Ask Federal Count to review SIFI Designation (15 January 2015), https://www.metlife.com/about/press-room/index.html?compID=155136; GE, Press Release, GE to Sell Most Capital Assets, Embrace Its Industrial Core (10 April 2015), http://www.gereports.com/post/116017450895/ge-to-sell-most-ge-capital-assets-embrace-its#.

[99] http://financialresearch.gov/about; http://www.gereports.com/post/116017450895/ge-to-sell-most-ge-capital-assets-embrace-its#.

do so by identifying core business lines and critical operations, funding, liquidity needs, interconnections, and interdependencies. Reporting entities include:

– Bank holding companies, including foreign banks with US operations and with $50 billion or more in total assets.
– Non-bank financial companies designated by FSOC as systematically important.

Title II—Orderly Liquidation Authority:

- In the USA, banks that have become insolvent traditionally are subject to a specific resolution regime under the Federal Deposit Insurance Corporation (FDIC), an approach that is distinct from the procedures generally applied to stock corporations under the US bankruptcy code. Title II has provided an alternative to bankruptcy by creating a systematic process to liquidate a non-bank financial institution or insurance company that is on the brink of failing. The FDIC is appointed receiver to liquidate the company.[100] The rationale behind this change in the law was, in part, an attempt to eliminate the market's presumption that certain entities are "too big to fail" and, by extension, to remove the need for future government bailouts.
- To determine if a financial institution should be placed in receivership under Title II, the Secretary of the Treasury considers two factors: (a) whether the company is in default, or in danger of default, and (b) the systemic risk involved in the potential default of the financial company. If, after due diligence, the FDIC believes it should be appointed as receiver, it will take control of the assets, obligations, and operations of the company.[101]
- Once it becomes the receiver, the FDIC assumes the powers required for an efficient liquidation. These include, but are not limited to, transferring or selling assets, creating bridge financial organizations that can help assume assets or liabilities, and approving valid claims against the company that will need to be paid. Title II also created the Orderly Liquidation Fund, a fund that will allow the FDIC to borrow funds from the US Treasury to cover the administrative costs of liquidation.

Title III—Consolidation of Bank Supervision: Transfer of Powers to the Comptroller of the Currency, FDIC, and Federal Reserve.

- Title III consolidated bank supervision in the USA to mitigate inconsistent enforcement of banking regulation among the various banking regulators. The title eliminated the Office of Thrift Supervision (OTS) as a standalone agency, and reassigned its duties to the Federal Reserve, OCC, and FDIC. The Fed is the regulator for the holding companies; the OCC is the depository institution regulator for federally registered banks, and the FDIC is the depository institution

[100] http://www.llsdc.org/assets/DoddFrankdocs/crs-r41350.pdf.
[101] Ibid.

regulator for state registered banks.[102] Three of the largest entities supervised by the OTS after failing during the financial crisis—Washington Mutual, IndyMac, and AIG—cumulatively represented a large part of the overall regulated portfolio of the OTS.

Title IV—Regulation of Advisers to Hedge Funds and Others:

- Title IV eliminates the "private investment adviser" exemption of the Investment Advisers Act. Private funds (e.g., hedge funds) with a certain amount of assets under management are now required to register with the SEC. Smaller advisors no longer fall under the SEC's jurisdiction and are monitored instead by their own state regulators.[103]

Title VI—Improvements to Regulation of Bank and Savings Association Holding Companies and Depository Institutions:

- Title VI amended the Bank Holding Company Act, establishing higher regulation for bank holding companies, savings and loan companies and depository institutions. In addition, it introduced the "Volcker Rule" that prohibits banking entities from engaging in proprietary trading.
- The Volcker Rule, named for former Federal Reserve Chairman Paul Volcker, prohibits an insured depository institution and its affiliates from engaging in certain activities including:
  - Proprietary trading;
  - The acquisition or retention of any equity, partnership, or other ownership interest in a hedge fund or private equity fund; and
  - Possessing an ownership interest in or sponsoring a hedge fund or private equity fund.[104]
- Board-supervised non-bank financial companies are subjected to additional capital requirements and quantitative limits if they engage in proprietary trading or maintain an ownership interest in, or sponsor, a hedge fund or a private equity fund.[105]

Collins Amendment:

- The Collins Amendment requires federal banking agencies to establish the minimum capital leverage and risk-based capital requirements to apply to insured depository institutions, bank and thrift holding companies and systemically important nonbank financial companies.[106]

---

[102] Dodd-Frank ActDodd-Frank Act § 312.

[103] Dodd-Frank Act § 403.

[104] http://www.davispolk.com/…/070910_Financial_Reform_Summary.pdf, Dodd-Frank Act § 619.

[105] Dodd-Frank Act § 619.

[106] Id. at 47.

- The minimum leverage capital and risk-based capital requirements applicable to these institutions are subject to two floors. They must be:
  - Not less than the generally applicable risk-based capital requirements and the generally applicable leverage capital requirements.
  - Not quantitatively lower than the above requirements that were in effect for insured depository institutions as of the date of enactment of the bill.

Title VII—Wall Street Transparency and Accountability

- Title VII created the framework for the regulation of the over the counter swap market. In particular, the SEC was tasked with regulating the security-based swap (SBS) market and the CFTC was given regulatory authority over all other swaps, such as energy and agricultural swaps. The Act mandated centralized clearing and exchange trading for many OTC derivatives. The clearing of swaps that the regulators determined should be cleared was required, and the ACT also implemented new mandatory trade execution requirements. Title VII introduced new types of market participants (swap dealers and major swap participants)[107] and swap trading and processing facilities (swap execution facilities, designated contract markets)[108] as well as expanding the definitions of certain market participants (future commission merchants, commodity pool operators)[109] and clearing organizations (derivatives clearing organizations).[110] These entities are all subject to registration and reporting requirements, and to oversight by the applicable regulator.

Title VIII—Payment, Clearing, and Settlement Supervision:

- Title VIII authorizes the FSOC to designate a Financial Market Utility (FMU) as "systemically important" if it determines that the failure of, or a disruption to, the functioning of the FMU could increase the risk of significant liquidity or credit problems spreading among financial institutions or markets, thereby threatening US financial stability.[111] In July 2012, the FSOC designated eight entities as FMUs.
  - The Clearing House Payments Company L.L.C.:
  - CLS Bank International
  - Chicago Mercantile Exchange
  - The Depository Trust Company
  - Fixed Income Clearing Corporation
  - ICE Clear Credit LLC
  - National Securities Clearing Corporation
  - The Options Clearing Corporation[112]

---

[107] Dodd-Frank Act § 730.

[108] Dodd-Frank Act § 733.

[109] Dodd-Frank Act § 724.

[110] Dodd-Frank Act § 723.

[111] Dodd-Frank Act § 804.

[112] FSOC Annual Report 2012, Appendix A: Designation of Systemically Important Financial Market Utilities, http://www.treasury.gov/initiatives/fsoc/Documents/2012%20Appendix%20A%20Designation%20of%20Systemically%20Important%20Market%20Utilities.pdf.

- The SEC, CFTC, and financial institution regulators are provided regulatory authority over the financial market utilities and organizations engaged in the payment, clearing, and settlement activities that they supervise. But Title VIII does not itself prescribe much detail as to the nature of such regulation.

Title IX—Investor Protections and Improvements to the Regulation of Securities:

- Title IX addressed a number of policy issues including increasing consumer protection; executive compensation and corporate governance; improvements to the regulation of credit rating agencies; improvements to the asset-backed securitization process (the process of turning mortgages, credit card loans, and other debt into marketable securities); and increasing regulatory enforcement and remedies.[113]

Title X—Bureau of Consumer Financial Protection:

- Title X established a Bureau of Consumer Financial Protection (CFPB) that has authority over consumer financial products including, among others, mortgages, credit cards, real estate settlement, money transmitting, and loan servicing.[114]
- The Bureau has the authority to administer, enforce, and implement federal consumer financial laws, which includes the power to make rules, issue orders, and issue guidance.
- One notable exception that can undermine the competence of the CFPB is that it is specifically precluded from exercising authority over entities regulated by the CFTC, SEC, or state securities regulators.[115]

Title XI—Federal Reserve System Provisions:

- The Federal Reserve engaged in emergency lending to non-bank financial firms as a result of the financial crisis of 2007 and 2008. Title XI amends the Federal Reserve Act, granting the Federal Reserve Board the ability to establish policies and procedures for emergency lending. The Act also allows the Government Accountability Office (GAO) to audit the Federal Reserve's lending facilities.[116]

Title XIII—Pay it Back Act

- Title XIII amends the stimulus measures by the US Government in attempting to stem the impact of the financial crisis under the Emergency Economic Stabilization Act of 2008 (EESA). The authority of the Secretary of the Treasury to purchase troubled assets under the Troubled Asset Relief Program (TARP) is reduced from $700 billion to $475 billion.[117] The proceeds of the sale of troubled assets purchased under TARP (Fannie Mae, Freddie Mac and Federal Home

---

113 http://www.llsdc.org/assets/DoddFrankdocs/crs-r41350.pdf.

114 http://www.llsdc.org/assets/DoddFrankdocs/crs-r41350.pdf.

115 Dodd-Frank Act § 1027(h), (i), (j).

116 http://www.llsdc.org/assets/DoddFrankdocs/crs-r41350.pdf at 5.

117 Dodd-Frank Act § 1302.

Loan Bank debt), any fees collected relating to those assets, and certain unutilized federal funds can only be used for deficit reduction.

Title XIV—Mortgage Reform and Anti-Predatory Lending Act.

- Title XIV established new requirements in the context of residential mortgage lending, including provisions applicable to loan origination, appraisal activities, and mortgage servicing. These changes were motivated in large part to address perceived abuses that were made apparent by the Global Financial Crisis, and to be administered by the CFPB established pursuant to Title X.

There are a few notable provisions of the Dodd-Frank Act on international coordination and harmonization. Section 175 (International Policy Coordination) provides generally that the FSOC, Secretary of the Treasury, Board of Governors of the Federal Reserve, as well as other government officials should consult with foreign financial regulatory authorities or international organizations on matters relating to systemic risk. This effort also encourages comprehensive and robust supervision, as well as regulation to protect financial stability. The newly created Federal Insurance Office within the Treasury Department is responsible for representing the USA in the International Association of Insurance Supervisors under Section 502. However, by far the most detailed references to international coordination and harmonization appear within the Title VII provisions on the regulation of swaps.[118]

The passage of the Dodd-Frank Act was itself a milestone. However, its implementation largely depends upon the promulgation of regulations by the myriad agencies authorized under its various titles, with the number of different rules estimated to total 390. As of this writing, over 5 years after the Act was passed, of the 390 total rulemaking requirements, 267, or slightly more than two-thirds, had been satisfied with finalized rules; proposed rules would meet 40, or 10 % more. As of this writing, rules had not yet been proposed to satisfy the remaining 83, or 21 %, of the rulemaking requirements.[119] Notably, the CFTC, compared to other agencies including the SEC, has relatively promptly implemented a number of major rules in its expanded areas of competence.

9.3.6 Regulatory Compliance and Enforcement

The macro-level approaches of the G-20, the FSB, and the SSBs on the implementation of financial standards have also been accompanied by action at the micro level to strengthen the focus of individual supervisors on effective implementation by the financial institutions they oversee. A perception that financial market rules were not always respected, or were not consistently applied, further undermined confidence

[118] See, e.g., Dodd-Frank Act § 752.

[119] See Davis Polk, Dodd-Frank Progress Report (Fourth Quarter 2015) at 2, available at http://www.davispolk.com/Dodd-Frank-Rulemaking-Progress-Report/.

in the financial sector. An actual (or perceived) lack of enforcement, or inconsistent enforcement, would undermine its intended effect regardless of how well drafted a requirement might be. Popular sentiment in the wake of the global financial crisis also called for individuals to be held responsible for the consequences of excesses or abuses in financial market activity.

In their initial 2008 summit, the G20 leaders, as part of their action plan for financial reform, agreed on the following in their efforts to promote the integrity of the financial markets: "*National and regional authorities should also review business conduct rules to protect markets and investors, especially against market manipulation and fraud and strengthen their cross-border cooperation to protect the international financial system from illicit actors. In case of misconduct, there should be an appropriate sanctions regime.*"[120]

At the EU level, the aforementioned Larosière report, outlining the plan for the European System of Financial Supervision, recognized sufficiently convergent sanctioning regimes as a necessary corollary to the new supervisory system: "*Supervision cannot be effective with weak, highly variant sanctioning regimes. It is essential that within the EU and elsewhere, all supervisors are able to deploy sanctioning regimes that are sufficiently convergent, strict, resulting in deterrence.*" [121]

The predecessors of the ESAs subsequently conducted a study of sanctions applied by the member states for violations of national rules, subsequently implementing some of the most important EU directives in the financial services area. These include MiFID and the directive implementing the BCBS capital requirements.

The study showed broad divergence and relevant weakness in the sanctions regime actually applied across the Member States. The EU Commission viewed this as risking the effectiveness of the financial service rules, with a potential negative impact on financial supervision.[122]

The Commission suggested that minimum common standards be established across the member states. These included enhanced administrative penalties, more significant monetary fines based in part on the benefit obtained through the violation, publication of the penalties, holding individuals as well as institutions accountable, and the possibility of criminal sanctions for the most significant violations.[123]

In comparison, the US supervisory authorities have had comparatively more authority to impose penalties upon financial services actors in cases of violations. Indeed, and perhaps more importantly, these authorities have been more regularly utilized than in other jurisdictions. There is a long history of significant penalties having been made public; in so doing the penalties have had a reputational impact as well as a direct pecuniary one. The foregoing notwithstanding, in the wake of the

[120] See G20 Washington Declaration, https://g20.org/wp-content/uploads/2014/12/Washington_Declaration_0.pdf (emphasis added).

[121] De Larosière, 2009: par. 201.

[122] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "Reinforcing sanctioning regimes in the financial services sector," COM (2010) 716 final (8 December 2010), pp. 9–10.

[123] Id. at 12–14.

global financial crisis there were many public calls for a more aggressive effort to hold institutions and individuals accountable for violations.

Many of the largest financial penalties globally—reaching hundreds of millions or even billions of dollars—were imposed by US authorities (both supervisors and criminal prosecutors) on global financial firms. Their US affiliates were also hit with financial sanctions for violations of the laws pertaining to market integrity, i.e., to anti-money laundering and financial sanctions. Other penalty actions were imposed for improprieties in the US mortgage markets that were exposed through the crisis.[124] Furthermore, penalties were imposed as part of efforts to combat abuse in connection with extraordinary stimulus and funding measures initiated in response to the crisis.[125]

From the perspective of this chapter's focus on the international capital markets, perhaps the most notable enforcement actions and resulting changes in business and regulatory compliance activities have been in benchmark indices. In July 2013, IOSCO issued its final report on Principles for Financial Benchmarks.[126] This work was undertaken as a direct response to allegations of a manipulation of interest rate benchmark indices such as the London Interbank Offer Rate (LIBOR). LIBOR was the basis for determining interest rates on a broad range of financial product from wholesale loans to retail products, such as residential mortgages and student loans, and to financial derivatives. While Benchmark Regulation is already in the rule-making process in Europe,[127] corresponding regulation is not planned in the USA.

9.4 The Future of Regulation in the EU and the USA

9.4.1 The European Union

The proposed establishment of a Capital Markets Union (CMU) in Europe will dominate the political agenda of the current European legislation from 2015 onwards. The broad set of measures as outlined in the CMU will significantly alter the regulatory environment for market participants and market infrastructure providers.

---

[124] See, e.g., US Department of Justice, Press Release, "Federal Government and State Attorneys General Reach $25 Billion Agreement with Five Largest Mortgage Servicers to Address Mortgage Loan Servicing and Foreclosure Abuses" (9 February 2012), http://www.justice.gov/opa/pr/federal-government-and-state-attorneys-general-reach-25-billion-agreement-five-largest.

[125] The Office of the Special Inspector General for the Troubled Asset Relief Program (SIGTARP), a sophisticated, white-collar law enforcement agency, was established by Congress in 2008 to prevent fraud, waste, and abuse linked to the $700 billion Troubled Asset Relief Program (TARP). See www.sigtarp.gov.

[126] Available at http://www.iosco.org/library/pubdocs/pdf/IOSCOPD415.pdf.

[127] See: http://ec.europa.eu/finance/securities/benchmarks/index_en.htm.

The European Commission's proposal can be regarded as the next step in the development and integration of European capital markets following the monetary and banking union. What started as a buzzword on the political agenda in mid-July 2014, when Jean-Claude Juncker made his opening statement to the EU Parliament, soon evolved into a comprehensive conceptual framework.

The CMU—as laid out by the European Union, the executive body of the European Commission—is, in summary, a joint vision for policy makers, as well as the industry and societal stakeholders, to:

- Further integrate and deepen European financial markets
- Strengthen Europe's competitiveness
- Increase its attractiveness for investors and companies

The main objective of the CMU is to achieve a more efficient allocation of capital throughout the EU. It will seek to accomplish this by developing non-bank sources of funding to foster economic growth and innovation, and to drive employment opportunities across Europe. In this sense, it is also a sign that the European political agenda has started to shift from crisis solving to fostering meaningful growth. And, hopefully, to closing the chapter of recession and anemic growth in Europe.

In an introductory Green Paper elaborating on the CMU concept,[128] the European Commission has identified five core principles as requirements to developing a reasonable CMU, namely (p. 5):

- Maximizing the benefits of capital markets for growth and jobs
- Creating a single market for all 28 member states, removing barriers to cross-border investment within the EU
- Effectively enforcing financial stability with a single rulebook for financial services
- Guaranteeing an effective level of consumer and investor protection
- Attracting investments from all over the world, and enhancing EU competitiveness

One of the most urgent needs for capital markets in general, and for achieving substantial non-bank funding reform in particular, is the revival of investor trust. This is true for companies as well as for individuals in order to restore the demand for new sources of funding. Measures for gaining trust should include:

- Communications: The promotion of the use of fair, efficient, and transparent markets that operate under highest possible standards.
- Education: Initiatives aimed at giving the broader public a greater understanding of the functioning of capital markets within the financial system, and knowledge of the benefits of attractive economics, which can be achieved through non-bank financing.

[128] The paper can be accessed online at http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=COM:2015:63:FIN&from=EN.

European markets are at a crossroads. The European system, unlike that of the USA, depends heavily on bank funding and liquidity: In the USA, only 30 % of general financing needs are covered by bank loans; the remaining 70 % is obtained via the capital markets. It is exactly the opposite in the EU.[129] In the wake of the financial crisis, intensified banking regulation has forced banks to strengthen their balance sheets, and bank funding has been decreasing in response to higher capital and liquidity requirements. That, in turn, has interrupted the highly essential provision of liquidity to the markets and, therefore, to the broader economy.

The EU, as a result, must attract companies of varying sizes to enter into equity and debt financing via primary markets. At the same time, their access to liquidity through secondary markets must be improved, this being an additional financing channel. Funding aside, companies require capital markets for hedging, and for minimizing the risks that arise from price fluctuations.

The CMU, however, should not be regarded as anti-banking. On the contrary, it practically defines a new role for banks as an essential pillar to support growth in Europe while guaranteeing a high level of consumer protection. In this context, banks will benefit significantly from the harmonization the CMU will bring to European markets. The benefits will be realized from higher transaction rates across the entire Union, from risks diversification and from increased intermediary services.

Ideally, bank funding and non-bank funding exist in parallel. Together, they provide companies with a range of choices for financing their essential investments.

A second pillar of the CMU concept is transparency since this both improves the quality of price discovery and reduces investment risk. Price transparency across all asset classes is crucial for investors in a stable financial market. And supervisors need additional data to spot potential market risks.

In reality, transparency is demanded by the general public along with some individual market participants while other market participants prefer not to lose their information advantage by disclosing trading data. Indeed, the rise of over-the-counter instruments and trading reflects this preference. Regardless, a lack of transparency can cause a serious threat to market stability, especially during stressful periods (as was evident in the latest financial crisis).

In this context, the advantage of integrated exchange organizations is that they are already operating fully regulated markets, replete with transparent order books and neutral price discovery processes. In this way, the exchanges are ideally equipped to support this principle, and to guarantee the highest level of investor and system protection. Hence, the CMU can be regarded as an opportunity for these institutions to offer new business possibilities, and as a means of supporting the development of reasonable regulation.

Finally, the CMU is committed to harmonization, as this will likely help to dismantle some of the cross-border barriers that retard the development of integrated European markets. Regulatory initiatives, especially those recently developed or reviewed, already honor the concept of harmonization. Nevertheless, significant

---

[129] Cicero analysis [4]: Unblocking the EU's capital markets.

fragmentation still exists in the public domain. The initiatives in this area, by building on a single rulebook that provides a harmonized regulatory framework, should significantly increase the attractiveness of European financial markets. This will enhance returns on investment and stimulate greater growth.

The EU, in its Green Paper,[130] defined both short- and long-term measures for all principles. This document also touched upon new regulatory initiatives, and on exiting initiatives considered to be subsumed under the CMU. From a market infrastructure provider's viewpoint, a functioning CMU should reconcile loose ends from finalization and implementation to the application of existing regulatory initiatives. Additionally, it must reduce the regulatory burden to create conditions that guarantee an attractive environment for companies and investors. It should also facilitate an efficient supervisory structure, and a global level playing field. Ultimately, this will better ensure European competitiveness in a fully globalized capital market.

Against this backdrop, the persistent call of some European countries for a Financial Transaction Tax (FTT) can be seen to be counterproductive: Neither the intended regulatory objectives, nor the financial goals, seem to be achievable with the introduction of a FTT. The tax, because it would increase trading costs, is deemed to have a negative effect on market-based financing. An additional negative effect is that the tax will likely be passed along to companies and savers, which will further decrease their willingness to invest in capital markets. The net effect will be to further erode investor trust and hinder the competitiveness of Europe.

Still, there are hopeful signs and improvements: The financial crisis revealed several weaknesses and shortcomings of International financial markets. Some European countries, in particular, are still struggling, as previously noted, with anemic growth rates in the aftermath. And the crisis saw a wave of regulatory initiatives flooding European financial markets that are aimed at solving immediate problems and stabilising the system. An extensive reform of financial markets that are significantly changing market structure followed with MiFID and EMIR coming to the fore. And, with markets and economies now recovering, the European regulators and supervisors have adjusted their focus to the final step on the road: stimulating growth and jobs.

As long as no unexpected events upend the financial markets in Europe or internationally, the coming years will be characterised, not by extensive reforms like MiFID or EMIR, but by the consolidation, review, amendment, or replacement of current regulatory reforms. The future of European financial markets lies in the deeper integration and specified initiatives that will foster growth and jobs throughout the European Union. First and foremost, for the financial markets infrastructure providers, this presents opportunities on various levels:

1. Regulators will rely on the experience and know-how of industry professionals when reforming financial market regulation in light of new requirements,

[130] The paper can be accessed online at http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=COM:2015:63:FIN&from=EN.

This provides the possibility to actively shape the debate as well as the resulting initiatives in a way that serves regulators, clients, and the industry equally.
2. The integration of financial markets in Europe fosters competition and provides opportunities to conquer new markets, and to introduce new products and services that serve the requirement of jobs and growth throughout the Union.

9.4.2 The USA

In the USA, instead of more change, the medium-term outlook of the financial regulatory environment is expected to be characterized by a period of consolidation, in particular in further implementation of the Dodd-Frank Act. Indeed, the pendulum has swing back from more regulation in the wake of the financial crisis, to increasing calls for regulatory relief as the US economy experienced slow but nonetheless generally positive growth.

From the moment the Dodd-Frank Act was passed, opponents and critics were calling for certain components of the legislation to be reversed. The most prominent categories of issues debated, and those most likely to be the subject of discussions for years to come, do not include the provisions of Title VII's OTC swaps, or the general principles of financial stability regulation, notwithstanding efforts of individual firms to fall under such regulation. Rather, the contested issues have focused in particular on the restrictions under the Volcker Rule, on the regulation of the mortgage markets, and on the scope of the CFPB.

Regarding market structure, in January 2015, the SEC announced the establishment of a new Equity Market Structure Advisory Committee to inform the SEC on the structure and operations of the US equities markets.[131] The Committee's inaugural meeting in May 2015 focused on Rule 611 of Regulation NMS, the Order Protection Rule, and in particular criticisms that Rule 611:

- Contributed to excessive fragmentation among trading venues
- Indirectly led to more dark pool trading[132] by narrowing the nature of competition on lit venues to factors such as speed, fees, and exotic order types, at the expense of factors such as liquidity and stability that are of major importance to investors
- Harmed institutional investors that need to trade in large size by forcing them to access small-sized quotations and, thereby, to signal their trading intentions to other, largely short-term proprietary traders

---

[131] See SEC, Equity Market Structure Advisory Committee, https://www.sec.gov/spotlight/equity-market-structure-advisory-committee.shtml.

[132] C.f. Chap. 2.

- Has not succeeded in achieving the SEC's stated objective of enhancing the reward for the display of limit orders[133]

SEC Chairman, Mary Jo White, has placed a priority on enhancing the US equity market structure, and so further regulatory proposals in this area can be expected.

Hence, when it comes to financial markets, the coming years promise opportunities for all market participants to contribute to stable, transparent, efficient, and prosperous markets for Europe and the international community.

References

1. Atlantic Council (2013): *The Danger of Divergence. Transatlantic Financial Reform & the G20 Agenda:* http://www.atlanticcouncil.org/publications/reports/the-danger-of-divergence-transatlantic-financial-reform-the-g20-agenda
2. BIS (2008): *Addressing financial system procyclicality. A possible framework*: http://www.financialstabilityboard.org/wp-content/uploads/r_0904e.pdf?page_moved=1
3. Blommestein, H. J.; Keskinler, A.; Lucas, C. (2011): *Outlook for the Securitisation Market.* OECD Journal: Financial Market Trends, Volume 2010, Issue 1: http://www.oecd.org/finance/financial-markets/48620405.pdf
4. Cicero Group (2015): Unblocking EU's Capital Markets. The European Commission's Green Paper on the Capital Markets Union: http://www.cicero-group.com/wp-content/uploads/2015/02/CMU-Green-Paper.pdf
5. CPSS-IOSCO (2012): *Principles for Financial Market Infrastructures:* http://www.bis.org/cpmi/publ/d101a.pdf
6. Eichengreen, B. (2011): *Reforming the International Financial Architecture*: http://eml.berkeley.edu/~eichengr/Reforming_Intl_Arch_2011_Ed.pdf
7. European Commission on MiFID: http://ec.europa.eu/finance/securities/isd/mifid2/index_de.htm
8. European Commission Press Release announcing MiFID II enactment (12 June 2014): http://europa.eu/rapid/midday-express-12-06-2014.htm?locale=en
9. Farrel, D. (2008): *New Thinking for a New Financial Order.* Harvard Business Review, September 2008 Issue: https://hbr.org/2008/09/new-thinking-for-a-new-financial-order.
10. FSB (2009): *Report of the Financial Stability Forum on Addressing Procyclicality in the Financial System*: http://www.financialstabilityboard.org/wp-content/uploads/r_0904a.pdf.
11. G20 (2009): *Leaders' Statement. The Pittburgh Summit, September 24-25 2009:* http://www.bundesregierung.de/Content/DE/StatischeSeiten/Breg/G8G20/Anlagen/G20-erklaerung-pittsburgh-2009-en.pdf?__blob=publicationFile&v=2
12. G20 (2014): 2014 *Agenda For Financial Regulation*: https://www.g20.org/g20_priorities/g20_2014_agenda/financial_regulation
13. Hüpkes, E. (2012): *The Architechture of International Financial Regulation*. In: Blair, M; Walker, G.; Willey, S. (Eds.): Financial Markets and Exchanges Law. Oxford University Press, Oxford.
14. IMF et al. (2009): *Report to G20 Finance Ministers and Governors. Guidance to Assess the Systemic Importance of Financial Institutions, Markets and Instruments: Initial Considerations*: http://www.bis.org/publ/othp07.pdf.

[133] See SEC Division of Trading and Markets, Memorandum to the Markets Structure Advisory Committee re Rule 611 of Regulation NMS (30 April 2015), available at https://www.sec.gov/spotlight/emsac/memo-rule-611-regulation-nms.pdf.

15. Jobst, A. (2008): *Back to Basics. What is Securitization?* Finance and Development, September 2008 Issue: http://www.imf.org/external/pubs/ft/fandd/2008/09/pdf/basics.pdf.
16. Larosière, J. de (2009): *The High-Level Group on Financial Supervision in the EU. Report:* http://ec.europa.eu/finance/general-policy/docs/de_larosiere_report_en.pdf
17. Minsky, H. P. (1992): *The Financial Instability Hypothesis*, Levy Economics Institute, Annandale on Hudson: http://www.levyinstitute.org/pubs/wp74.pdf
18. Shiller, R. J. (2003): *The New Financial Order. Risk in the 21st Century*. Princeton University Press, Princeton: http://digamo.free.fr/shiller2003.pdf
19. Willey, S. (2012): *The European System of Financial Supervision.* In: Blair, M; Walker, G.; Willey, S. (Eds.): Financial Markets and Exchanges Law. Oxford University Press, Oxford.

Further Reading

Gomber, P.; Nassauer, F. (2014): Neuordnung der Finanzmärkte in Europa durch MiFID II/MiFIR. Safe White Paper Series No. 20, Frankfurt am Main: http://safe-frankfurt.de/uploads/media/Gomber_Nassauer_MiFIDII_MiFIR.pdf

Steil, B. (2013): *The Battle of Bretton Woods*, Princeton University Press, Princeton.

Official EU Documents

COM(2011) 206 final: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0206:FIN:en:PDF

COM(2015) 63 final: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=COM:2015:63:FIN&from=EN

Consolidated Version of the Treaty on the Functioning of the European Union: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT&from=EN

Directive2004/39/EC:http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:02004L0039-20110104&from=DE

Directive 2013/36/EU: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32013L0036&from=EN

Directive 2014/65/EU: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=OJ:L:2014:173:FULL&from=EN

Dodd-Frank Wall Street Reform and Consumer Protection Act – H. R. 4173 (111th): http://www.cftc.gov/ucm/groups/public/@swaps/documents/file/hr4173_enrolledbill.pdf

ESRB Regulation: https://www.esrb.europa.eu/shared/pdf/ESRBen.pdf?5020438634e48ca0076187a2a62e9344

EuGH Joined Cases C-261/07 and C-299/07 on Directive Directive 2005/29/EC (Unfair Commercial Practice"): https://webgate.ec.europa.eu/ucp/public/index.cfm?event=public.guidance.showArticle&elemID=52

EU Benchmark Regulation: http://ec.europa.eu/finance/securities/benchmarks/index_en.htm

European Union Legal Acts: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=URISERV:ai0032&from=EN

Glass-Steagall-Act – H.R. 5661 (73rd): https://fraser.stlouisfed.org/scribd/?title_id=991&filepath=/docs/historical/congressional/1933_bankingact_publiclaw66.pdf

Regulation (EC) No 764/2008: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:218:0021:0029:en:PDF

Regulation (EU) 600/2014: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=OJ:L:2014:173:FULL&from=EN
Regulation (EU) No 1092/2010: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32010R1092&from=EN
Regulation (EU) No 1093/2010: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32010R1093&from=EN
Regulation (EU) No 1094/2010: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32010R1094&from=EN
Regulation (EU) No 1095/2010 : http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32010R1095&from=EN
Regulation (EU) No 575/2013: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32013R0575&from=EN
Regulation (EU) No 648/2012:http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32012R0648&from=DE
Regulation (EU) No 909/2014: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32014R0909&from=EN
SEC(2011) 1227 final: http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:52011SC1227&from=DE,
United States-European Union Financial Markets Regulatory Dialogue Joint Statement of Jan 2015: http://ec.europa.eu/finance/general-policy/docs/global/150115-us-eu-joint-statement_en.pdf
WpHG online: http://www.gesetze-im-internet.de/wphg/__33a.html

Part II
Trading of Financial Instruments: Selected Topics

Chapter 10
Security Market Microstructure: The Analysis of a Non-frictionless Market

Reto Francioni, Sonali Hazarika, Martin Reck, and Robert A. Schwartz

10.1 Introduction

Security market microstructure addresses issues that involve the implementation of portfolio (investment) decisions in a marketplace. Implementation entails the placement and handling of orders in a securities market, and their translation into trades and transaction prices. The process links fundamental information concerning equity valuation (which is of primary concern to portfolio managers) to prices and transaction volumes that are realized in the marketplace. The quality of the link depends on the rules, procedures, and facilities of a securities market, and on the broader regulatory and competitive environment within which the market operates.

Widespread interest on the part of the securities industry, government, and academia is testimony to the importance of market microstructure analysis. The subject addresses issues that concern investors, broker/dealer intermediaries, market regulators, exchanges, and other trading venues as well as the broad economy. Interest in microstructure has increased sharply over the past three and a half decades, spurred in particular by three events: the US Securities and Exchange Commission's (SEC) Institutional Investor Report (1971), the passage by the US Congress of the Securities Acts Amendments of 1975, and the sharp stock market drop on October 19, 1987. Further, the advent of computer-driven trading in recent years has enabled researchers to capture electronically the full record of all trades and quotes, and this has provided empirical researchers with far richer data (referred to as "high-frequency data") for analyzing trading and price setting.

This chapter includes material from [1] and from Schwartz which was reprinted in [2].

R. Francioni (✉) • M. Reck
University of Basel/Deutsche Börse Group, Basel, Switzerland
e-mail: reto.francioni@unibas.ch

S. Hazarika • R.A. Schwartz
Zicklin School of Business, Baruch College, CUNY, New York, NY, USA

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_10

Over the years, microstructure analysis has expanded and, concomitantly, exchange structure has strengthened. We consider both of these developments in this chapter. First, we set forth the major challenges that the microstructure literature addresses. Second, we consider the properties of a frictionless trading environment. Third, we present a broad view of the direction in which microstructure analysis has been and is evolving. Fourth, we turn to one application—the design of an actual marketplace: Deutsche Börse's electronic trading system, Xetra. The German market was the last of the major European bourses to introduce an electronic trading platform, and it is state of the arts, which makes Deutsche Börse a particularly interesting case in point. Fifth, in the concluding section, we consider the bumpy and hazardous road that takes us from theory to the development of an actual marketplace.

10.2 Microstructure's Challenge

Microstructure analysis has four broad applications. First (and this is a key focus of the chapter), it gives guidance to market structure development. The link with market structure is straightforward: the critical factor that drives microstructure analysis is friction in the marketplace (i.e., the explicit and implicit costs of implementing portfolio decisions), and trading costs depend on the architecture of the marketplace which determines how orders are handled and turned into trades. The flipside of friction is illiquidity, and a primary function of a market center is to amass liquidity.

Microstructure's second application is to facilitate the development of trading strategies and algorithms for asset managers and broker/dealer intermediaries. The importance of this application is evident in the current development of computer-driven algorithmic trading. Algorithms can be fine-tuned to take account of, for example, the probability of a limit order executing, time-of-day effects such as market openings and closings, search for liquidity in a fragmented environment, and choice of a trading modality (e.g., a continuous limit order book market, a quote-driven dealer market, a periodic call auction, a block trading facility, or hybrid combinations of the above).

The third application of microstructure analysis concerns tests of market efficiency. In the 1970s, at a time when the subject was first emerging, the efficient market hypothesis (EMH) was widely accepted by financial economists as a cornerstone of modern portfolio theory, and it continues to receive broad academic support today. The hypothesis addresses informational as distinct from operational efficiency (the latter refers to the containment of transaction costs by superior market design). According to the EMH, a market is informationally efficient if no participant is able to achieve excess risk-adjusted returns by trading on currently available information. Many of the EMH tests have considered one major part of the information set—market information (e.g., recent quotes, trading volume, and transaction prices). If prices properly reflect all known information, then (in a frictionless market at least) they must change randomly over time, hence the term "random walk." Earlier studies, based on daily data, generally supported the random walk hypothesis. However, with the advent of high-frequency data, the footprints of complex correlation patterns have been detected. This observation, along with superior knowledge of the impact of trading costs on returns behavior, is casting a new light on market efficiency. Whether inefficiency is thought of in operational or informational terms, the EMH is not as stellar as it once was.

In its fourth application, microstructure analysis sheds light on how new information is incorporated into security prices. In a zero-cost, frictionless environment, share values would be continuously and instantaneously updated with the release of new information. In actual markets, however, information must be received and assessed, traders' orders must be placed and processed, and executions must be delivered and accounts cleared and settled. Costs, both explicit (e.g., commissions) and implicit (e.g., market impact), are incurred throughout this chain of events. Highlighted in much microstructure literature are the costs that some participants incur when, in an asymmetric information environment, other participants receive information first and trade on it to the disadvantage of the uninformed.

Asymmetric information is not the only reality, however. In light of the size, complexity, and imprecision of much publicly available information, one might expect that investors in possession of the same (large) information set will form different expectations about future risk and return configurations. This situation is referred to as "divergent expectations."[1] Asymmetric information and divergent expectations together reflect a rich set of forces that impact the dynamic behavior of security prices.

This overview of microstructure's four broad applications underscores that trading frictions are the subject's raison d'être. Participant orders cannot be translated into trades at zero cost (markets are not perfectly liquid), and trades typically are not made at market clearing (i.e., equilibrium) prices. Trading decision rules (algorithms) are needed because the costs of implementing portfolio decisions can sharply lower portfolio performance. In fact, much algorithmic trading is designed to control trading costs, rather than to exploit profitable trading opportunities. Today, trading is recognized as an activity that is both distinct from investing and equivalently professional. Market structure is of concern to the buy-side desks precisely because markets are not perfectly liquid, and neither are they perfectly efficient, either informationally or operationally. Consequently, better market structure can deliver superior portfolio performance for participants.

What is the economic service, one might ask, that an equities market provides? The fuzzy link that connects information and prices in the non-frictionless environment underscores two major market functions—price discovery and quantity discovery. Price discovery refers to participants collectively searching for equilibrium prices. Quantity discovery refers to the difficulty that participants who would be willing to trade with each other actually have finding each other and trading when markets are fragmented. This difficulty is accentuated because some participants (primarily institutional investors) do not immediately reveal the total size of their orders (doing so would unduly drive up their market impact costs).

[1] For a recent discussion, see Davis et al. [72].

Market structure affects both the accuracy of price discovery and the completeness of quantity discovery. The link between market structure and price discovery depends on the environment within which participants are operating. At one end of the spectrum, investors can be equally informed and form homogeneous expectations based on the information. At the other end, they can be differentially informed and form divergent expectations with regard to commonly shared information. When investors who share common information all agree on share values (i.e., have homogeneous expectations), prices can be "discovered" in the upstairs offices of research analysts. When investors are not equally informed, and when they form different expectations based on common information, prices must be discovered in the marketplace. In this second environment, the economic service provided by an exchange is clear—it "produces the price."

Regarding quantity discovery, handling the orders of large institutional customers is a challenge. It is not at all uncommon for an institution to want to buy or to sell, for instance, 500,000 shares of a company that has an average daily trading volume of 300,000 shares. Executing an order of this size can easily drive prices away from the trader before the job has been completed. The adverse price move is a *market impact cost*. Institutions attempt to control their market impact costs by trading patiently and, as much as possible, invisibly. Good market structure can help. To this end, a number of alternative trading systems (ATSs) have been formed in recent years, and dark (i.e., non-transparent) liquidity pools have emerged.

With prices discovered in the marketplace, participants employ trading strategies when they come to the market to implement their portfolio decisions. Participants with differential information that will soon become public determine how best to meter their orders into the market so as to move prices to new levels with *minimal* speed. Additional questions that any trader might ask include the following: "If I trade now, at the current moment, how will the price that I will receive compare with the average price that shares are trading at today?" "Is price currently at a sustainable, validated level, or is it likely to move higher or lower in the coming hours, minutes, or even seconds?" "Would I do better to be patient and place a limit order, or submit a market order and get the job done right away?" "Should I attempt to trade now in the continuous market, or wait for a closing call?" The orders that a set of participants reveal to the market depend on how questions such as these are answered, and prices that are set and trading volumes that are realized depend on the orders that are revealed.

The categories of trading costs that receive the most attention on the part of exchanges, regulators, and academicians are generally those that are the most straightforward to measure: commissions and bid-ask spreads. Increasingly precise measures of market impact are also becoming available, and this cost too is being widely taken into account. On the other hand, the opportunity cost of a missed trade, being far more difficult to quantify, is often overlooked. Also more challenging is quantifying a cost that has received little formal attention: realizing executions at poorly discovered prices. The problem, of course, is that equilibrium values are not observable and appropriate benchmark values are not easily defined.

10.3 The Perfectly Liquid Environment of CAPM

Peter Bernstein's [3] piece in the *Journal of Portfolio Management* has the intriguing title, "The Surprising Bond Between CAPM and the Meaning of Liquidity." In it he wrote, "The more liquid an asset, the greater the dominance of systematic risk over stock specific risk." We build on this insight in this section. In so doing, we formalize the fact that the capital asset pricing model describes an extreme case, a totally frictionless world where liquidity is infinite and systematic risk has complete dominance over stock-specific risk. The analysis provides a good platform from which to launch a discussion of market microstructure, the study of a non-frictionless environment.

CAPM models the price of the individual equity shares that, in aggregate, comprise the market portfolio. Following standard methodology, we start our analysis of the frictionless environment by taking the market portfolio to be one single asset (e.g., an all-encompassing exchange traded fund). We consider the demand of an agent to hold shares of this one risky asset when the only alternative is the riskless asset. We show that an individual agent's demand curve to hold shares of the risky asset is downward sloping, and then use this curve to re-derive certain key CAPM equations to show that the associated demand to hold shares of each individual equity issue in that portfolio is infinitely elastic, and that therefore the market for the individual shares is infinitely liquid.

In the CAPM world, each individual equity issue in the market portfolio has an intrinsic value that is given by the parameter that locates the height (on the price vector) of that infinitely elastic demand. In the section that follows, we turn to the non-frictionless environment of microstructure analysis where individual stock demand curves are downward sloping, the liquidity of individual shares is, therefore, finite, and individual shares do not have intrinsic values.

To obtain the representative investor's demand curve to hold shares of the risky market portfolio, first we state the agent's utility (of wealth) function. The demand curve to hold shares of the market portfolio may then be obtained directly from the utility function. The derivation follows Ho et al. [4].

We make the following assumptions:

- The investor's portfolio comprises a risk-free asset and one risky asset (shares of the market portfolio).
- Share price and share holdings are continuous variables.
- Short selling is unrestricted.
- The existence of a brief trading period, T_0 to T_1, which is followed by a single investment period, T_1 to T_2.
- All transactions made during the trading period are settled at point in time T_1.
- The investor seeks a portfolio at the beginning of the investment period (at time T_1) that will maximize the expected utility of wealth to be realized at the end of the investment period (at time T_2).

- Investor expectations with respect to the share price at the end of the investment period (at time T_2) are exogenously determined (expectations are independent of the current price of shares).
- Investors are risk averse.

The following variables are used:

C_0 = holdings of the risk-free asset at the beginning of the trading period (T_0).
C_1 = holdings of the risk-free asset at the beginning of the investment period (T_1).
N_0 = number of shares of the market portfolio held at the beginning of the trading period (T_0).
N_1 = number of shares of the market portfolio held at the beginning of the investment period (T_1).
$R_0 - 1$ = risk-free rate of interest over the trading period.
$R_1 - 1$ = risk-free rate of interest over the investment period.
P_1 = price at which shares of the market portfolio are purchased or sold during the trading period.
P_2 = price at which shares of the market portfolio can be sold at the end of the investment period (T_2).
$r_m = P_2/P_1 - 1$ = return on the market portfolio.
Q = number of shares traded by the investor at the beginning of the investment period (T_1); $Q > 0$ indicates a purchase; $Q < 0$ indicates a sale.

10.3.1 The Expected Utility of End-of-Period Wealth

The participant starts the *investment period* with C_1 dollars of the risk-free asset and N_1 shares of the market portfolio (the risky asset). Therefore, wealth at T_2 is given by $C_1R_1 + N_1P_2$. As of T_1, this wealth is uncertain because P_2 is uncertain. As of T_1, the expected utility of end of period wealth can be written as

$$EU\left(C_1R_1 + N_1P_2\right) \tag{10.1}$$

The investor starts the *trading period* with C_0 dollars of the risk-free asset and N_0 shares of the risky asset. If during the trading period the decision maker were to exchange holdings of the risk-free asset for Q shares of the risky asset at a price of P_1, the *expected utility of end-of-period wealth*, written as a function of P and Q, given N_0 and C_0, would be

$$h\left(P_1, Q \mid N_0, C_0\right) = EU\left[\left(C_0R_0 - QP_1\right)R_1 + \left(N_0 + Q\right)P_2\right] \tag{10.2}$$

where $C_0R_0 - QP_1 = C_1$ and $N_0 + Q = N_1$. Equation (10.2) can be rewritten as

$$h\left(P_1, Q \mid N_0, C_0\right) = c + gQ\left(a - bQ - P_1\right)$$

where

$$
\begin{aligned}
c &= U(W) - \pi N_0^2 U'(W) / R_1 \\
g &= U'(W) R_1 \\
a &= \left[E(P_2) - 2\pi N_0\right] / R_1 \\
b &= \pi / R_1 \\
\pi &= -\frac{1}{2}\left[U''(W) / U'(W)\right] \mathrm{Var}(P)
\end{aligned}
\tag{10.3}
$$

The step from (10.2) to (10.3) involves a Taylor expansion of the investor's utility around the expected value of wealth if the investor does not trade.[2] The procedure is a convenient way of introducing the variance term into the utility function.[3]

10.3.2 The Reservation Demand Curve

Equation (10.3) can be further assessed with the use of risk aversion and risk premium measures that are defined in Appendix 1. Specifically, using (10.3), we now obtain both a *reservation price demand curve* and an *ordinary demand curve* to hold shares of the risky asset. We consider the reservation demand curve first.

The reservation price for a purchase or a sale is the maximum price the decision maker would be willing to pay to buy a given number of shares ($Q>0$), or the minimum price the decision maker would be willing to receive to sell a given number of shares ($Q<0$) when the only alternative is not to trade at all. Equation (10.3) shows that, if no trade is made (that is, if $Q=0$), the decision maker's expected utility is equal to c. The reservation price for any value of Q is the price that equates the expected utility $[h(P_1, Q|N_0, C_0)]$ if the trade is made, with the expected utility (c) if no trade is made. Thus the reservation price for any value of Q is given by

$$
h\left(P^R, Q \mid N_0, C_0\right) = c \tag{10.4}
$$

where P^R is the reservation price associated with the trade of Q shares. Given (10.3), for (10.4) to be satisfied, we must have $a - bQ - P_1 = 0$. Hence the reservation price demand curve is

$$
P^R = a - bQ \tag{10.5}
$$

[2] For a discussion of the Taylor procedure see, for example, R. G. D. Allen, *Mathematical Analysis of Economists,* London, England: Macmillan, 1960.

[3] Two further assumptions are required to obtain (10.3): (1) the third derivative of utility with respect to wealth is small enough to ignore; and (2) the squared deviation of the expected rate of return on the risky asset from the risk-free rate is small enough to ignore.

10.3.3 The Ordinary Demand Curve

Using (10.3), we can also obtain the ordinary demand curve. At any value of P_1, the decision maker selects the value of Q that maximizes expected utility. Hence, the ordinary price demand curve is given by

$$\frac{\partial h}{\partial Q}\left(P^0, Q \mid N_0, C_0\right) = 0 \tag{10.6}$$

where P^0 is the "ordinary" price associated with the trade of Q shares. Differentiating h in (10.3) with respect to Q, setting the derivative equal to zero, and rearranging gives

$$P^0 = a - 2bQ \tag{10.7}$$

10.3.4 The Risk Premium and the Market Price of Risk

When the investor has traded the optimal number of shares of the market portfolio at the market determined price per share, his or her risk premium can be related to the market price of risk. Assessing, the ordinary demand curve at $P^0 = P_1$ gives

$$P_1 = \frac{E(P_2)}{R_1} - \frac{2\pi N_1}{R_1} \tag{10.8}$$

Multiplying by R_1/P_1, rearranging, and recognizing that $[E(P_2)/P_1] - 1 = E(r_m)$ and $R_1 - 1 = r_f$, we get

$$\frac{2\pi N_1}{P_1} = E(r_m) - r_f \tag{10.9}$$

Therefore, we have

$$\pi_{M\%} = E(r_m) - r_f \tag{10.10}$$

where $\pi_{M\%}$ is the marginal risk premium (see Appendix 1). Note that the right-hand side is the price of risk. We thus see that the investor achieves an optimal holding of the risky asset by obtaining the number of shares that equates the marginal risk premium with the market price of risk.

10.3.5 The Investor's Optimal Point on the Capital Market Line

The demand model can be used to assess the investor's optimal point on the capital market line. Let r_p be the return on the combined portfolio (N_1 shares of the market portfolio and C_1 dollars of the risk-free asset). From Appendix equation (10.28) we have

$$\pi = \pi_{rp}\left(\frac{W}{N_1}\right)^2$$

which, using $R_A = -U''(W)/U'(W)$, the measure of absolute risk aversion, can be written as

$$\pi = \frac{1}{2}R_A Var\left(r_p\right)\left(\frac{W}{N_1}\right)^2 \tag{10.11}$$

Because $\sigma_p = (NP/W)\sigma_m$, we have $\mathrm{Var}(r_p) = \sigma_p(NP/W)\sigma_m$ and can write (10.11) as

$$\pi = \frac{1}{2}R_A\sigma_p\left(\frac{PW}{N_1}\right)\sigma_m \tag{10.12}$$

Substituting (10.12) into (10.9) and simplifying give

$$R_R\sigma_P = \frac{E\left(r_m\right) - r_f}{\sigma_m} \tag{10.13}$$

where $R_R (= WR_A)$ is the measure of relative risk aversion.

Equation (10.13) shows that for the investor to hold an optimal combined portfolio, the market price of risk per standard deviation of the market portfolio must be equal to the investor's coefficient of relative risk aversion times the standard deviation of the combined portfolio's return.

Letting $w = N_1P_1/W$, substituting $w\sigma_m, = \sigma_p$ into (10.13), and rearranging give

$$w = \frac{E\left(r_m\right) - r_f}{Var\left(r_m\right)R_R} \tag{10.14}$$

Equation (10.14) shows that the percentage of wealth that the risk-averse participant invests in the market portfolio is positively related to the expected return $E(r_m)$, and negatively related to r_f, $\mathrm{Var}(r_m)$, and R_R. Investors all face the same values of $E(r_m)$, r_f, and $\mathrm{Var}(r_m)$, but differ according to their degree of risk aversion. More risk-averse investors (larger R_R) have smaller optimal values of w and hence are more apt to lend at the risk-free rate (which implies $w<1$); less risk-averse investors (smaller R_R) have larger optimal values of w and hence are more likely to borrow at the risk-free rate (which implies $w>1$).

The right-hand side of (10.13) is the market price of risk per standard deviation of the market portfolio. The total compensation for risk taking is the price of risk times the standard deviation that the investor accepts (here, the standard deviation of the combined portfolio). Multiplying both sides of (10.13) by σ_p, we obtain

$$R_R\, Var\left(r_p\right)=\left[\frac{E\left(r_m\right)-r_f}{\sigma_m}\right]\sigma_p \tag{10.15}$$

Adding r_f to both sides of (10.15) gives the investor's total compensation for waiting and for risk taking:

$$E\left(r_p\right)=r_f+R_R\, Var\left(r_p\right)=r_f+\left[\frac{E\left(r_m\right)-r_f}{\sigma_m}\right]\sigma_p \tag{10.16}$$

Equation (10.16) shows that the location of the investor's optimal point on CAPM's capital market line depends on his or her measure of relative risk aversion (R_R).

10.3.6 The i*th Risky Asset's Point on the Security Market Line*

We now assess the demand model to show the location of an ith risky asset on the security market line. In so doing, we establish that the demand for the ith risky asset is infinitely elastic. Equation (10.10) shows that the marginal risk premium for each investor, as a percentage of P_1, will equal $E(r_m)-r_f$. Therefore, for each investor,

$$\frac{R_A\, Var\left(P_2\right)N_1}{P_1}=E\left(r_m\right)-r_f \tag{10.17}$$

It follows from Equation (10.17) that investors with lower values of R_A hold a larger number of shares, such that the product R_AN_1 is the same for all investors. Because $r_m=(P_2/P_1)-1$, $Var\left(r_m\right)=Var\left(P_2\right)/P_1^2$. Substituting $Var\left(r_m\right)P_1^2=Var\left(P_2\right)$ into (10.17) and simplifying give

$$R_A\, Var\left(r_m\right)P_1N_1=E\left(r_m\right)-r_f \tag{10.18}$$

Using $P_1N_1=wW$ we obtain

$$wR_R\, Var\left(r_m\right)=E\left(r_m\right)-r_f \tag{10.19}$$

Equation (10.19) can be interpreted as an equilibrium condition for each investor. Because $wR_R = R_A N_1 P_1$, and given that the product $R_A N_1$ is constant across investors, $R_R w$ is constant across all investors. [It is also clear from (10.19) that the product wR_R must be constant across all investors, because $E(r_m)$, r_f, and $\mathrm{Var}(r_m)$ are the same for all.]

The equilibrium condition for each investor with respect to the market portfolio implies an equilibrium condition for each investor with respect to any ith risky asset in the market portfolio. The CAPM shows that the relevant measure of risk for the ith risky asset is $\beta_i = \sigma_{im}/\mathrm{Var}(r_m)$. Therefore, writing $\mathrm{Var}(r_m) = \sigma_{im}/\beta_i$, substituting into (10.19), and multiplying both sides by β_i we get

$$wR_R\sigma_m = \beta_i\left[E\left(r_m\right) - r_f\right] \tag{10.20}$$

Adding r_f to both sides of (10.20) gives CAPM's security market line,

$$r_f + wR_R\sigma_{im} = r_f + \beta_i\left[E\left(r_m\right) - r_f\right] = E\left(r_i\right) \tag{10.21}$$

where $E(r_i)$ is the expected return on the ith stock in the market portfolio. Equation (10.21), assessed at $w = 1$, shows that the expected return for the ith risky asset depends on its covariance with the market return, and on the measure of relative risk aversion for an investor whose optimal combined portfolio contains the market portfolio only. The equation also shows that the ith risky asset's specific location on the security market line depends on the covariance of the asset's return with the return on the market portfolio, and hence that its expected return depends only on β_i, its systematic risk.

It follows from the above discussion that the demand to hold shares of the market portfolio is downward sloping, while the demand for each individual stock in the market portfolio is infinitely elastic. The reason is that perfect substitutes do not exist for the aggregate portfolio, but they do exist for the individual stocks. Only one factor characterizes any ith stock—β_i, its covariance with the market. But the covariance for any stock can be duplicated exactly by an appropriate combination of two or more other stocks, and all holdings that have the same covariance must yield the same expected return. If they were to yield different expected returns, an unlimited number of shares of the higher yielding position would be bought, and an unlimited number of shares of the lower yielding position would be sold short until, with costless trading, the buying and selling pressures bring the two prices into exact equality. Unlimited buying (selling) at any price lower (higher) than the beta appropriate, CAPM price manifests an infinitely elastic demand to hold shares. That is, at an infinitesimally higher price no shares will be held, and at an infinitesimally lower price demand will be unlimited.

Bernstein's [3] two insights immediately follow: a stock's systematic risk totally dominates its specific risk, and the market for each ith stock is infinitely liquid at the price which translates into $E(r_i)$, its systematic risk-appropriate return.

As we turn to the non-frictionless market, the infinitely liquid, infinitely elastic property of CAPM is a good point of departure from the frictionless world. A common denominator in many microstructure analyses is that the demand to hold shares of individual stocks is downward sloping (which means that shares do not have intrinsic values). Market makers post bid and ask quotes which, when raised, result in more public sales to the market maker and which, when lowered, result in more public purchases from the market maker. Bid and ask quotes can be distributed over multiple price points in competitive dealer markets as well as on public limit order books. Trading is not costless. Both explicit costs (e.g., commissions and taxes) and implicit costs (e.g., market impact costs) are incurred. Information is complex and imprecise, and thus investors commonly disagree about its interpretation. Arbitrage is not costless, and perfect substitutes for individual issues do not exist. Share values depend not only on the calculations of systematic risk in the upstairs markets, but also on how orders interact in the marketplace. As a consequence of all of this, trades that are made and the transaction prices that they are made at also depend on the structure of a marketplace. Microstructure analyses address these realities that CAPM does not comprehend.

10.4 What Microstructure Analysis Has to Offer: Personal Reflections

In this section we review the development of microstructure analysis. Our objective is not to provide a comprehensive survey of the literature, but to highlight some of the important themes that have given guidance to market structure development. More detailed information can be obtained from Cohen et al. [5] who have provided an early survey of the field; from O'Hara [6] who discusses important theoretical microstructure models; from Madhavan and Ananth. Market Microstructure. Journal of Financial Markets [7]; Biais et al. [8]; and Parlour and Seppi [9] who have provided more recent surveys; and from Hasbrouck, Joel. Empirical Market Microstructure. Oxford University Press [10] who deals with empirical microstructure research and research methodology. We first focus on the early literature, next turn to more recent developments, and lastly present our thoughts concerning an important direction in which future microstructure research ought to head.

10.4.1 The Early Focus

The first contributions to the new field in financial economics that came to be called "microstructure" were made by a couple of people who participated in the SEC's Institutional Investor Report (1971). A handful of others independently started to focus on microstructure topics in the early 1970s. Eventually a few of the early researchers came to recognize the commonality of their interests and, applying the title of Garman's [11] well-known paper, "Market Microstructure," they gave the field its name.

Much of the early literature focused on dealers and exchange specialists. These market makers were viewed as the suppliers of immediacy to investors, and the spread was considered the price they charge for providing this service in an environment where order arrival is nonsynchronous. Of key importance was the relationship between spreads and the costs of market making.

The earlier market maker studies were in large part motivated by a desire to determine whether or not these intermediaries were realizing monopoly profits and, if so, whether or not their profits were attributable to market making being a natural monopoly. Spreads that are greater than the costs of market making would be taken as an indication of monopoly power on the part of the dealers, and spreads that were negatively related to trading volumes would indicate economies of scale in market making, which could imply a natural monopoly [12]. Spreads were indeed found to decrease with transaction volume, but reasons other than market making being a natural monopoly were advanced [13, 14].

The general picture which emerged was that the trading costs incurred by investors could be lowered by strengthening competition between market maker intermediaries. In particular, competition in the NYSE market was deemed inadequate, as specialists and the exchange itself were viewed as having monopoly positions: each stock was assigned to just one specialist; the NYSE's order consolidation rule (Rule 390) precluded in-house executions by requiring that exchange members send their orders for NYSE-listed securities to an exchange; and commissions were fixed and unjustifiably high [15].[4]

Not surprisingly, the focus on the market maker firms led several researchers to model market maker pricing decisions (i.e., the setting of their bid and ask quotes). These included Bagehot et al. [16], Stoll et al. [17], Amihud et al. [18], Ho et al. [19–21], and Mildenstein et al. [22]. With one exception [16], the early formulations dealt with inventory considerations. A market maker firm holding an undesirably long position would lower the quotes (i.e., lower the offer so as to sell more shares, and reduce the bid so as to discourage others from selling shares to it). Reciprocally, a market maker who was short would raise the quotes. This response on the part of the public (buy more shares when the market maker's offer is lower, and sell more share when the market maker's bid is higher) is evidence that the public's demand to hold shares of any specific stock was taken to be downward sloping. A variety of mathematical tools were used to solve for optimal market maker quotes. These models also gave further insight into the cost components of the market maker's spread [23].

While insightful, the early inventory-based pricing models suffered from some shortcomings. First, the early formulations for the most part assumed monopoly market makers, even though some of these models were applied to markets such

[4] Another major issue addressed by the microstructure literature at that time was the impact of information on trading volume and price ([74–76]).

as the New York Stock Exchange where exchange specialists were in fact competing with other floor brokers and customer limit orders [24]. The application of theory further suffered from the reality that the price of immediacy for an investor is not the spread of an individual market maker, or even an average market maker spread, but the inside spread (i.e., the lowest ask across all market makers minus the highest bid).[5] It is important to note that dealer spreads could individually remain relatively invariant with respect to transaction volume while the inside spread fell appreciably.

A further shortcoming of most of these earlier models is that they did not take account of a major cost incurred by market makers: the losses generated by trading with better informed investors. Recognition of this reality (which is also outside the scope of the frictionless world of CAPM) led to a development that did much to establish microstructure as an important new field in financial economics—the introduction of market maker models that were based, not on inventory management, but on controlling the cost incurred when some investors are in possession of information that the market maker and other investors have not yet received. Bagehot et al. [16] was the first to embark on this line of thought. He was later followed by, among others, Glosten et al. [25] and Kyle, Albert. "Continuous Auctions and Insider Trading." Econometrica, 53 [26].

With information asymmetries, the market maker always looses when trading with a better informed participant. For microstructure theorists at the time, this meant that, for the dealer market not to fail, some investors must trade for reasons that are not related to information.[6] Liquidity considerations (i.e., an investor's personal cash flow needs) were one such motive for public buying and selling. A third participant type was also introduced along with the liquidity traders—noise traders (participants who trade on price moves as if they contain information when in fact they do not). This trio of informed traders, liquidity traders, and noise traders was used to show how markets could function and, in so doing, enable new information to be incorporated into security prices (Grossman et al. [27], Milgrom, Paul and Nancy Stokey. "Information et al. [28]," Kyle, Albert. "Continuous Auctions and Insider Trading." Econometrica, 53 [26], Glosten et al. [25], Copeland et al. [29], and Easley and Maureen O'Hara. "Order Form and Information in Securities Markets." Journal of Finance 46 [30–32]).

At this stage in its early development, the microstructure pricing models were predominantly market maker models. One exception should be noted, however: a National Book System proposed by Mendelson et al. [33] contained a comprehensive description of an order-driven automated trading system that provided guidance for designing the first exchange-based electronic trading systems. For a more recent discussion of automated trading systems, see Domowitz, Ian and Benn Steil. "Automation et al. [34]." Most equity markets around the globe are now order-driven, limit order

[5] For further discussion, see Cohen et al. [5].

[6] A market supported by informational trading only can indeed function if agents trade with each other because their expectations are divergent. When the information that triggers trading is common knowledge, the condition may be thought of as one where agents are agreeing to disagree.

book markets that might include market makers in a hybrid structure (as does the NYSE), but are not basically quote-driven (i.e., dealer) markets (as was the old Nasdaq and London Stock Exchange). The limit order book markets are driven by the orders placed by the investors themselves, not by market maker intermediaries.

10.4.2 The Current Focus

Over the years, microstructure analysis has grown extensively on both the theoretical and empirical fronts. Concomitantly, the securities markets themselves have evolved, becoming evermore technologically developed, more global in outreach, but also more fragmented between different trading facilities. One important new direction microstructure research has taken is to further model the order-driven market, an environment where natural buyers and sellers provide immediacy to each other because some, who are patient, are willing to post limit orders while others, who demand immediacy, choose to submit market orders that execute against the posted limit orders. Understanding the costs of, and motives for, placing limit orders as distinct from market orders was called for.

With limit orders, the very existence of the bid-ask spread has to be explained. That is, with a sufficiently large number of participants placing priced orders, one might expect that orders would be posted at virtually every available price point in the neighborhood of equilibrium, and that the spread would disappear. Cohen, Maier, Schwartz, and Whitcomb (CMSW) made this point in their review paper [5], and they analyzed the existence of the spread in Cohen et al. [35].[7] They further write, “With regard to modeling the market spread, we suggest that a straightforward aggregation from individual spreads is not possible in a system where there is no clear distinction between demanders and suppliers of immediacy, and where traders meet in a dynamic, interactive environment that incorporates the impact of investor order placement strategies.” Strategic order placement clearly required further analysis.

The task, however, was not simple. Some of the first papers in this area assumed, as is true for a dealer market, that limit order and market order participants are two separate, exogenously fixed groups that are separated by a firewall [36]. This assumption, while simplifying mathematical modeling, unfortunately distills out much of the richness of an order-driven market. More recent models have eliminated the firewall (Handa et al. [37]; Foucault et al. [38]; Parlour et al. [39]; Handa et al. [40]; Foucault et al. [41]; and Goettler et al. [42]). With the choice between limit order and market order endogenous, for any market to function, participants must divide naturally into four groups which reflect two dichotomies (one between buyers and sellers and the other between limit order and market order placers), not the standard two (buyers and sellers).

[7] Cohen et al. [35] describe the trade-off between execution probability and price improvement in the optimal choice between limit and market orders.

With order type selection endogenous in the order-driven market, the balance between immediacy demanders and immediacy suppliers becomes a second equilibrium that must be understood. That is, one needs to recognize the conditions under which some participants will choose to be liquidity demanders (place market orders) while others choose to be liquidity suppliers (place limit orders). If a reasonable balance is not achieved between these two groups, the order-driven market will fail (as indeed it does for thinner, small cap stocks). Increasingly, these issues have been handled, and some sophisticated limit order models have been developed.[8]

Microstructure analysis of trading systems has expanded to include periodic call auctions.[9] The economics of a call auction are quite different from those of continuous trading and, consequently, so too are the order placement strategies that participants should employ when they approach a call market. Call auctions do not, by their very nature, supply immediacy. Rather, orders that are entered during a call's book-building phase are held for a periodic crossing at a single clearing price at the (generally predetermined) time of the market call. Consequently, buy and sell orders submitted to a call do not execute when they arrive even if they match or cross in price (matching and crossing orders execute immediately in a continuous trading environment). This being the case, limit and market orders have a different meaning in a call: limit orders do not supply immediacy to market orders, and market orders are simply extremely aggressively priced limit orders (i.e., a market order to sell in a call effectively has a limit price of zero, and a market order to buy effectively has a limit price of infinity).

Today, virtually all modern, electronic exchanges open and close their continuous markets with call auctions. Consequently, participants face further decisions when operating in a call plus continuous, hybrid market: how to submit an order to a call auction which is followed by continuous trading (e.g., an opening call), and how to submit an order to a continuous trading environment that is followed by a call auction (e.g., a closing call). Taking these tactical decisions into account is part of the complexity of microstructure analysis.

Technological developments have simultaneously enabled new trading venues to emerge (which can fragment markets) while providing connectivity between them (which can consolidate markets). Concurrently, regulatory initiatives have been motivated by the desire to intensify inter-market competition. Questions can be raised, however, concerning fragmentation of the order flow. The conventional wisdom has been that the consolidation of order flow improves liquidity, and exposing each order to all other displayed orders gives investors the best prices for their trades. Consolidating trading in a single market provides incentives to liquidity suppliers to compete aggressively for market orders by revealing their trading interest, and by being the first to establish a more favorable price (if time is used as a secondary priority rule).

On the other hand, arguments in favor of trading on multiple markets include the benefits of inter-market competition, and the fact that traders with disparate motives for trading may want different marketplaces to trade in (i.e., the "one-size-does-not-fit-all" argument). And so, different markets develop to serve diverse investor needs

[8] See Back et al. [77] for a recent discussion and further references.

[9] See Economides et al. [78] for a description of alternative call market structures.

(such as achieving a faster execution vs. obtaining a better price). One growing need among large institutional investors, the ability to trade large orders with minimal market impact, has led to the advent of dark pool, block trading facilities such as Liquidnet, Pipeline, and ITG's Posit that aid in quantity discovery. This development in the industry has spawned a related line of research on off-exchange and upstairs trading ([43–46]).

A spectrum of market quality issues have been of long and continuing importance to microstructure researchers. These include market transparency,[10] both pre- and post-trade [47], the accentuation of intraday price volatility, and correlation patterns which have been observed in high-frequency data [48]. Other important issues include price clustering and tick sizes [49–51]. Applications such as transaction cost analysis (TCA) and algorithmic trading have received increasing attention [52]. The relative performance of floor-based vs. electronic trading is another important issue [34].

A major line of empirical research was pioneered by Hasbrouck et al. [53, 54] who decomposes transaction prices into two components: a random walk component and a stationary component. The random walk component is identified with an efficient price that the market is trying to discover. The stationary component is viewed as microstructure noise. Microstructure noise is commonly explained by features such as the bid-ask spread, market impact, and discreteness of the pricing grid. The noise component has also been attributable to price discovery itself being a dynamic process (Menkveld et al. [55], and Paroush et al. [56]).[11]

Numerous empirical studies have focused on two of the world's premier markets, the New York Stock Exchange and Nasdaq [54, 57–65], among others. Many other studies have considered European markets, Asian markets, and other markets around the world (e.g., [66–68]).[12] Across all of these markets, structural and performance differences have been noted, but also major similarities have been observed. It is apparent that, despite the influence of historic and cultural considerations, trader behavior and market performance around the globe depend largely on microstructure realities. Alternatively stated, trading rooms and markets around the world bear striking resemblances to each another.

Another recent line of research has considered how search costs affect bid-ask spreads in financial markets. To this end, Duffie, Pedersen, and Garleanu [69] present a dynamic model of market makers under the assumption of no inventory risk and information that is symmetrically distributed. They show that sophisticated investors who have better search and bargaining abilities face tighter bid-ask spreads. This is in contrast to traditional information-based models which imply that spreads are wider for more sophisticated (i.e., better informed) investors.

---

[10] Trading systems differ in their degree of transparency Pagano et al. [79] investigate whether greater transparency enhances market liquidity by reducing the opportunities for taking advantage of uninformed participants.

[11] Also see Hasbrouck, Joel. "One security et al. [65], Harvey et al. [80], and Jones et al. [81]." Further references are provided by Menkveld et al. [55].

[12] Also see Bessler, Wolfgang. Editor, Bösen et al. [82] for discussion and further references.

As we have noted, unlike in the frictionless market arena of CAPM, amassing liquidity is a primary function of a marketplace and market structure features are generally designed with liquidity implications in mind. Asset managers also take liquidity into account, along with the two other standard variables of modern portfolio theory, risk, and return. Difficulties in defining, measuring, and modeling liquidity are formidable, however, and the literature that deals with it directly is relatively sparse [70]. Nevertheless, liquidity considerations have permeated the microstructure literature, both explicitly and implicitly.[13]

Looking back over the development of microstructure analysis, two observations stand out. First, microstructure studies have in multiple ways given direction to market structure development. Second, to a remarkable extent, the various theoretical microstructure models that are center stage today, and many empirical analyses that are based upon them, share a common structural framework—the asymmetric information paradigm. This consistency is desirable in that it implies that the field has grown by accretion rather than by replacement. Consequently, new insights are more apt to refine than to contradict old conclusions.

Consistency, however, is not desirable if the common structural framework becomes overly rigid and restrictive, and if it yields incomplete and/or misleading answers to questions involving trader behavior, market structure, and regulatory policy. At times, a literature starts to advance along new fronts. We consider this possibility next for the microstructure literature.

10.4.3 Future Directions

As we have noted, the current focus in the literature is on asymmetric information-based models, which are characterized as follows. Trading is driven by informational change, liquidity needs, and noise trading. The information motive for trading is the first mover of the three (liquidity and noise trading are required so that a market will not fail). Further, order arrival in the continuous environment is generally taken to be asynchronous. For a continuous trading regime to function with asynchronous order arrival, the presence of a limit order book and/or a market maker intermediary is required.

Information trading is of keen interest because it represents the process by which new information is reflected in share values. In the standard asymmetric information models, it is assumed that all participants in possession of the same information form equivalent expectations concerning future risk and return configurations. When information changes, however, participants may not all receive the news at the same time; some receive it before others, a reality that, at any point in time, can divide traders into two groups—the informed and the uninformed.

[13] For further discussion and references regarding liquidity see Amihud et al. [83]; Chordia et al. [84, 85]; Hasbrouck, Joel and Duane Seppi. "Common Factors in Prices et al. [86]; Amihud et al. [87]; and Pástor et al. [88]."

Informed participants will never trade with each other; consequently, liquidity and noise traders must be present for a market to function. As noted, asymmetry of information, for the most part, lies at the heart of the standard microstructure models of today.

The homogeneous expectation assumption has been tempered of late. As a further departure from the infinitely liquid, zero-cost environment of CAPM, it is being recognized that some participants produce "private information" (namely, that they further process information so as to gain insights that are not immediately available to others). Whether participant expectations differ because of the actual production of private information, or simply because different people interpret the same information or news announcement differently, the expectations of a group of investors can be divergent.

Also at the heart of the asymmetric information models is the presumption that a stock has a fundamental value that bears a unique relationship, not to trader activity in the marketplace, but to the fundamental information that informed traders possess. The process of information being fully reflected in prices under asymmetric information is the act of informed and uninformed agents trading with each other until any discrepancy between a market price and a fundamental value is eliminated. The process can be viewed as arbitrage. In the earlier dealer models, the market maker was assumed to know a stock's fundamental value. In later models, informed traders but not the market maker know the fundamental values [26]. Especially in the later models, price discovery is not instantaneous; rather, it is a protracted process that depends on the individual strategies employed by the informed and uninformed agents.

In recent years, an alternative paradigm has been emerging: a divergent expectations environment [71]. While institutionally realistic, this paradigm has met with considerable academic resistance. For one thing, homogeneous expectations environments are far easier to deal with mathematically and homogeneity has, in many applications, proven to be a useful modeling assumption. The assumption has also been retained for another reason. As an attribute of individual rationality, it is presumed that intelligent agents facing the same information and applying the same (correct) analytic techniques will reach the same conclusions and, therefore, will have homogeneous expectations.

Fundamental information, however, is enormous in scope. It is complex and imprecise, and our tools for analyzing it are relatively crude. In the presence of fuzzy information, expectations can be divergent. Allowing for divergent expectations opens another path for microstructure analysis, and it introduces new questions concerning agent behavior, market structure, and regulatory policy. Moreover, a further element can enter the analysis in a divergent expectations environment: along with forming their own opinions, agents may also respond to the opinions of others, i.e., exhibit adaptive valuation behavior [56, 72].[14] Just how agents commu-

[14] Adaptive valuation behavior refers to individual agents becoming more bullish (bearish) when learning of the relatively bullish (bearish) attitudes of others.

nicate with each other and respond to each others' opinions is a subject for ongoing research. The topic also opens another interface with behavioral finance.

Price discovery acquires a different meaning in a divergent expectations environment, and this has important implications for market structure. When asymmetric information characterizes a community of investors, the strategic behavior of informed agents can affect the path that price takes when news moves a share value from one equilibrium to another, but the new equilibrium is path *independent*. With divergent expectations, the new equilibrium is path *dependent*—it depends on how the opinions of a diverse set of agents are integrated [56]. Alternatively stated, with divergent expectations, price discovery is a coordination process and, as such, is directly effected by market structure.

In the standard asymmetric information environment, the key dichotomy is between informed and uninformed participants. But a second dichotomy also exists—one that separates large institutional customers from small retail customers. One might expect that the informed investor set would largely comprise the institutional customers. After all, the institutions are professional, they can afford to continuously monitor information and respond to news, and their very size (all else constant) reduces their per share cost of doing so. With divergent expectations, however, there is no presumption that institutional customers can, because of their size, consistently evaluate shares more accurately. On the contrary, institutions commonly disagree with each other and, as a consequence, commonly trade with each other.

In the divergent expectations environment, institutional investors do not necessarily have an advantage over retail customers as fundamental analysts. In fact, their size makes trading more difficult and they incur higher transaction costs. So what accounts for their popularity? The value added by the mutual funds, pension funds, etc. comes largely from their ability to facilitate diversification. Further, they can bring a systematic, professional, and disciplined approach to portfolio management [72].

10.5 From Theory to Application

Microstructure analysis is inherently involved with analyzing the detailed functioning of a marketplace. The literature has a strong theoretical component and, to a large extent, is structured to yield insights into the effect of market design (structure and regulation) on market performance. Hopefully, theory can provide a broad roadmap for real-world market architects to follow. In this section we provide a broad overview of major technology and regulatory changes that have taken place in the USA and Europe.[15]

[15] Further discussion of market structure development is provided by Harris and Larry. Trading and Exchanges: Market Microstructure for Practitioners [89].

10.5.1 Technological Developments

Two exogenous forces have driven market structure change: technology and regulation. Regarding technology, the first big step was taken in 1971 in the USA when the National Association of Securities Dealers (NASD) introduced an electronic automated quotation (AQ) display system called NASDAQ. The Toronto Stock Exchange was the first exchange to introduce an electronic order-driven platform, its Computer Assisted Trading System (CATS); the year was 1977. Following in Toronto's footsteps, London instituted SEAQ in 1986, Paris rolled out its Cotation Assistée en Continu (CAC) in 1986, and Deutsche Börse's Xetra came to life in 1997. Also in 1997 the London Stock Exchange introduced its Stock Exchange Trading System (SETS) limit order platform. By the end of the twentieth century most of the exchanges in Europe had converted to electronic limit order book platforms.

Change came more slowly in the USA. Instinet introduced an electronic platform in 1969. Nearly 30 years later, Instinet became known as an Electronic Communications Network (ECN). In short order, a slew of other ECNs emerged, led most prominently by Archipelago and Island. In 2002, Nasdaq implemented its own electronic platform which, at the time, was called "SuperMontage." Most recently, in the Spring of 2006, the newly privatized NYSE Group initiated its Hybrid Market, a facility that has transformed the Big Board from a floor-based "slow" market into a hybrid that includes a "fast market" electronic venue. As of this writing, the floor-based component of the NYSE's hybrid has been markedly reduced in importance. Several specialist firms have ceased operations, other floor brokers have departed, and the trading room areas have collapsed from five to two.

10.5.2 Regulatory Initiatives

Major regulatory initiatives have played an important role in jump-starting these market structure changes. The 1975 Congressional Securities Acts Amendments was the first sizable regulatory foray into market structure development. The Amendments precluded the fixing of commission rates and mandated the development of a National Market System (NMS). In 1997, the US Securities and Exchange Commission instituted its new Order Handling Rules (OHRs), which require that market makers holding customer limit orders display those orders in their quotes, and that dealers at least match any quotes that they themselves display on an ECN (either by bettering the quotes that they offer customers or by posting their superior quotes in Nasdaq's SuperMontage). Following the OHRs, three other regulatory initiatives were introduced in the USA in relatively fast succession. In 2000, the NYSE, under pressure from the SEC, rescinded its order consolidation rule (Rule 390). In 2001, the US markets completed the transition from fractional to decimal pricing, which resulted in the minimum tick size decreasing from 1/16 of a dollar or 6.25

cents (it had earlier been 1/8 of a dollar or 12.5 cents) to one cent. In 2005, the SEC adopted Regulation NMS, the key provision of which is that better priced limit orders cannot be traded through (the trade-through rule was fully implemented in 2007).

On the eastern side of the Atlantic, the first major regulatory initiative was taken in 1993 when the Investment Services Directive opened the door for cross-border trading by introducing the single European passport. As discussed in Schwartz, Robert and Reto Francioni. Equity Markets in Action and Sons [2], "*Passporting* defines a system of mutual acceptance of other EU countries' rules without truly harmonizing all of the details of the various rules." Major regulatory change is currently coming again to the European arena in the form of the Markets in Financial Instruments Directive (MiFID). Key provisions in MiFID include a best execution requirement (echoes of the 1975 US Securities Acts Amendments), a quote disclosure requirement for upstairs broker/dealers (echoes of the US Order Handling Rules), and the disallowance of order focusing rules (echoes of the US SEC pressuring the withdrawal of NYSE Rule 390). A major regulatory difference is that no trade-through rule has been imposed on the European markets (unlike under the US SEC's Reg NMS).

10.6 Deutsche Börse: The Emergence of a Modern, Electronic Market

We turn in this section to the designing of an actual marketplace. Our focus is on Deutsche Börse: it is the dominant stock exchange in Germany, the last of the major European bourses to go electronic, and its technology is state of the art.

Important insights were gained from the microstructure literature during Xetra's planning period and the system's implementation has marked a huge step forward for Germany's equity markets. But our roadmap, which is undoubtedly incomplete today, was even more limited in the 1994–1997 years when Xetra was being designed. And, there is always the danger that the cartographer whose map is being used has some misconceptions (e.g., believes in the existence of the Northwest Passage).

10.6.1 The German Equities Market in the Mid-1990s

As recently as the mid-1990s, the German market had major structural defects that would undermine its competitiveness in the European arena. In recognition of this, Deutsche Börse, the newly founded exchange operator of the Frankfurter Wertpapierbörse (FWB), became the leading force for change.[16]

[16] FWB also owned the futures and options exchange Deutsche Termine Börse. After the 1997 merger with SOFFEX, DTB became Eurex.

In the mid-1990s, Frankfurt's trading floor was the major marketplace for German stocks, but the German market was badly fragmented. Kursmaklers, the equivalent of specialists, concentrated much of the liquidity in their order books. A primitive (by today's standards) electronic trading system, IBIS (which was owned by FWB), operated in parallel with the floor trading. IBIS's central component was an open limit order book that had hit and take functionality, but did not match orders automatically. The electronic system captured about 40 % of the trading volume in the 30 large-cap DAX stocks, but no link existed between IBIS and the floor. Seven other floor-based regional exchanges were also operating in Germany with technical infrastructures that were similar to those in Frankfurt. In total, the regionals at that time were attracting roughly 10 % of German exchange-based trading volume. Moreover, off-board trading has been (and still is) prevalent in Germany [73].

Transparency for floor trading (pre-trade transparency in particular) was low. Quotes were not distributed publicly (they were available on the floor only). Price priority between different trading venues was not enforced and orders executed in one market commonly traded through orders waiting to be executed in another market. Market manipulation and other abuses of power and position were believed to be rife on the old Frankfurt floor. Given the appreciable market fragmentation, poor transparency, imperfect inter-market linkages, and dubious floor behavior, transaction costs were high. Changes, both structural and regulatory, were called for. The result was the development of Xetra, an electronic order-driven trading system that comprises two principal modalities—a continuous order book platform and periodic single-price call auctions.[17]

10.6.2 Designing a New Trading System

Xetra's development started in 1994, and the system was launched in 1997.[18] Strong external forces also motivated this reengineering of Deutsche Börse's market structure: regulatory reform, soaring trading volumes, pan-European harmonization of the exchange industry, vibrant cross-border competition for order flow, and rising concerns of market participants about the future performance of Germany's financial markets.

Through Xetra's design stage, microstructure theory, even as it existed at the time, was an indispensable guide. This new field in financial economics, with its origin in issues concerning the competitive and architectural structure of an equity market, should have been able to give guidance to the development of an actual marketplace such as Xetra. To an extent, it has fulfilled its promise. The literature gave Deutsche Börse a broad roadmap, and it has highlighted underlying relationships and other important considerations that a market architect should be aware of.

[17] For further discussion and descriptions, see Francioni et al. [1].

[18] Appendix 2 provides details of Xetra's design.

Building the Xetra model involved specifying principles that the new market should implement, and the system's functionality also had to be defined. Most importantly, the new market system was to provide equal and decentralized access to all of its participants. Further, the system's functionality and the market information delivered to users (both pre- and post-trade) were to be the same for all traders. A trader's location should not matter. With this in mind, Deutsche Börse's fundamental architectural decision was to structure a hybrid market that included two major modalities—a continuous electronic order-driven platform, and periodic call auctions that were used primarily for market openings and closings.[19]

An absolutely critical attribute of an order-driven trading system is its ability, vis-à-vis its competitors, to win the battle for liquidity. Regarding this matter, the earlier microstructure literature has given some guidance, but liquidity is a complex attribute to deal with. As it is not easy to define and measure, liquidity has been very difficult to model and assess. However, as noted above, the measurement and analysis of liquidity are currently attracting considerably more attention in the microstructure literature.

Price discovery and transparency are two other issues for which the microstructure literature has provided valuable guidance. The architects at Deutsche Börse recognized that price discovery is a primary function of a market center, and their major reason for introducing the call auctions was to sharpen its accuracy, particularly at market openings and closings. Understanding that transparency is important while recognizing that it should not be excessive, the decision was made to disclose only the indicative clearing price (not the full book of orders) in the pre-call, book-building period.

Microstructure literature has given insights into the operations of the public limit order book for continuous trading. At the time, recognition was also emerging of periodic call auctions, a modality that was clearly differentiated from, but could effectively be used with, the continuous market. With regard to continuous trading, microstructure analyses of the use of limit and market orders and of the interaction between these two order types proved to be most valuable. However, a deeper understanding of the economics of an order-driven market now exists than was the case in the 1994–1997 period when Xetra was being designed.

Another important contribution of microstructure theory has been the classification of traders according to their needs for immediacy and their propensities to be either givers or takers of liquidity. The differentiation between informed and uninformed traders also proved to be valuable, particularly with respect to the market maker role that has been incorporated into Xetra. Specifically, market makers, referred to as "designated sponsors," were included to bolster liquidity provision for smaller cap stocks. A balance had to be achieved between the obligations imposed on the designated sponsors and the privileges granted to them. To accomplish this, information had to be assessed concerning the role of dealers in general (e.g., NASDAQ-type market makers) and specialists in particular (e.g., NYSE-type

---

[19] Interestingly, the microstructure literature on call auctions was relatively sparse at that time. For an early discussion, see Handa et al. [90].

specialists). That balance defined the designated sponsors' role in Xetra, and secured their willingness to accept it. Market microstructure insights also yielded the understanding needed to transform the specialist role into the newly designed designated sponsor role.

But designing an automated trading systems is indeed a complex task, and the gap between theory and implementation is both large and intricate. Trading decisions can be made in a large variety of ways that run the gamut from humans interacting directly with humans without computers to humans trading via electronic order handling and execution systems and to computers making trading decisions that are sent electronically to a computerized market (e.g., computer-driven algorithmic trading). Since the mid-1990s, market structure development has involved mainly the design of an electronic trading facility.

Deutsche Börse took account of the fact that automation impacts both the way in which trading decisions are made and the process by which prices are determined and trades executed in a market center. An electronic market requires the specification of an array of critical features (e.g., the trading modalities employed, rules of price and quantity determination, and basic features such as order types and trading parameters). With an electronic market, the software that implements a desired market structure must be specified on a level of detail that far exceeds what is required for human intermediated trading.

For instance, a human agent (specialist) has historically handled price determination at NYSE openings. This function is performed with reference to various rules, but the specialist is also free to exercise reasonable judgment. Further, human-to-human interactions can evolve naturally as problems, opportunities, and new competitive pressures arise. In contrast, with a fully electronic opening, every possible condition that can occur must be recognized and a rule for dealing with it specified, and electronic interaction can be changed only by rewriting the code that specifies with step-by-step precision just how orders are handled and turned into trades and transaction prices.

How does one achieve the precise specifications that a computerized trading system must have? In 1994, the market architects at Deutsche Börse could study the operations of other electronic platforms (e.g., CATS in Toronto and CAC in Paris). Doing so was helpful but of limited value given that Deutsche Börse was looking to develop a distinctive system.

When moving into new territory, market structure development is a venture. How does one know in advance whether or not it will work? How can one determine whether or not the new system will be viable from a business perspective? Nevertheless, design decisions have to be made, technical requirements must be specified, and the system must be built. The decisions involved represent huge financial bets on whether or not a new market structure will attract sufficient liquidity. Prototyping a new market in the design phase helps the assessment process, but doing so was considerably more difficult in 1994 than it is today with the advent of superior information technology and testing capabilities. In 1994, the architects were forced to rely more on their own educated judgment and on any insights they might gain from microstructure research.

Those who are involved in the design of an actual market realize that the devil is in the details. Consider, for instance, the specification of a call auction. A call has excellent theoretical properties, but how should an actual auction be designed? It is straightforward to say that the market clearing price in a call auction should be the value that maximizes the number of shares that trade. But what should the specific rule be for selecting the clearing price if two prices both result in the same maximum trade size? Additionally, how transparent should the book be in the pre-call, order entry period? Are further design features needed to counter the possibility of gaming? And so on.

Other considerations that for the most part are outside the scope of the microstructure literature also came into play during the design of Xetra. Information technology issues such as scalability, open architecture, and system reliability are of critical importance. So too are procedures for post-trade clearing and settlement. One of the final steps in the structural design of the new German market was the introduction in 2003 of a central counterparty (with a CCP, counterparty risk management was centralized and trading became fully anonymous, both pre- and post-trade). Electronic trading is also a prerequisite for highly efficient straight-through processing (STP involves all stages of a trade's life cycle). Information technology has further facilitated the timely capture of market data (all trades, quotes, market index values, etc.) and has expedited its delivery to users. With regard to these diverse applications, Deutsche Börse has achieved a closer integration between trading on Xetra and the broader market infrastructure.

10.7 Conclusion: The Roadmap and the Road

A market architect must have a roadmap that, broadly speaking, says where one ought to head and roughly how to get there. To this end, the microstructure literature has added clarity, articulation, and intellectual support. Briefly stated, the objective is to reduce trading frictions (costs), sharpen price discovery, and facilitate quantity discovery. The means of achieving this broad objective involve the amassing of liquidity. This is done through the appropriate use of limit order books for both continuous and call auction trading and, where appropriate, the inclusion of broker/dealer intermediaries. Further insights are gained from microstructure's in-depth analyses of trading motives (new information, liquidity needs, and technical trading signals). The literature has also provided guidance with regard to issues such as transparency and the consolidation (fragmentation) of order flow.

But theory, even if it does provide a good roadmap, can take one only so far. The closer one gets to the design of an actual system, the more apparent the complexities of trading and trading systems become. The road actually traveled is indeed bumpy and hazardous. System designers know that "the devil is in the details." They have to grapple with issues ranging from scalability, reliability, and other IT requirements to business considerations concerning the ultimate profitability of a trading venue. The market architects at Deutsche Börse recognized these issues and their new system, Xetra, has marked a huge step forward for the German equity market.

Today, important problems persist with regard to market design in Germany (and in all other markets around the world). Two fundamental questions concerning market architecture that have yet to be adequately answered are the following: (1) What is the best way to deal with large, institutional orders? (2) How is liquidity creation best handled for mid-cap and small-cap stock? At the same time, important microstructure topics continue to emerge at the academic research desks. Are there limits beyond which microstructure theory cannot provide guidance? Are there limits to the level of efficiency that a real-world market can ever achieve? Undoubtedly, both answers are "yes" but, without question, neither of these limits has as of yet been reached. Quite clearly, microstructure research and the design of an actual marketplace remain works in progress.

10.8 Appendix 1: Risk Aversion and Risk Premium Measures

Our analysis of the perfectly liquid CAPM environment makes reference to two measures of risk aversion and to several dimensions of a risk premium. We provide details concerning both of these in this appendix.

10.8.1 Risk Aversion

We use two risk aversion measures: (1) $R_A = -U''(W)/U'(W)$ is a measure of absolute risk aversion, and (2) $R_R = WR_A$ is a measure of relative risk aversion. Because $U'' < 0$ for a risk averse decision maker, R_A, $R_R > 0$ for risk aversion. Larger values of R_A and R_R indicate higher degrees of risk aversion. R_A is a measure of absolute risk aversion because it reflects the decision maker's reaction to uncertainty in relation to the *absolute* (dollar) gains/losses in an uncertain situation. R_R is a measure of relative risk aversion because it reflects the decision maker's reaction to uncertainty in relation to the *percentage* gains/losses in an uncertain situation.[20]

10.8.2 Risk Premiums

A *risk premium* is the minimum dollar compensation a decision maker requires to hold a risky asset in place of an alternative that involves no risk. Specifically, a decision maker would be indifferent between a riskless investment with a certain return of D dollars and a risky investment with an expected dollar return of $E(Z)$ equal to

[20] For further discussion, see J. Pratt, "Risk Aversion in the Small and the Large," *Econometrica*, January 1964.

D plus the investor's risk premium. In general, the investor's risk premium depends upon his or her utility function and initial wealth, and upon the distribution of Z.

Pi (π) in (10.3) is a risk premium: π equals one-half of R_A (the measure of the investor's absolute risk aversion) times Var(P_2), which measures the absolute (dollar) risk attributable to holding one share of the market portfolio. The uncertainty associated with holding N shares of the risky asset is Var(NP_2) = N^2 Var(P_2); thus the total risk premium for holding N shares is

$$\pi_T = \pi N_1^2 \tag{10.22}$$

Dividing (10.22) by N_1 (= $N_0 + Q$) gives the risk premium per share (the average risk premium):

$$\pi_A = \pi N_1 \tag{10.23}$$

Differentiating (10.22) with respect to N_1 gives the risk premium for a marginal share (the marginal risk premium):

$$\pi_m = 2\pi N_1 \tag{10.24}$$

Dividing (10.24) by P_1 expresses the marginal risk premium as a percentage of current price:

$$\pi_{M\%} = \frac{\pi_M}{P_1} = \frac{2\pi N_1}{P_1} \tag{10.25}$$

The return on the combined portfolio of N_1 shares of the market portfolio and C_1 dollars of the risk-free asset is

$$r_P = \left(\frac{P_2}{P_1} - 1\right)\left(\frac{P_1 N_1}{W}\right) + \left(1 - \frac{P_1 N_1}{W}\right) r_f \tag{10.26}$$

and the variance of the return on the combined portfolio is

$$Var\left[\left(\frac{P_2}{P_1}\right)\left(\frac{P_1 N_1}{W}\right)\right] = \left(\frac{N_1}{W}\right)^2 Var\left(P_2\right) \tag{10.27}$$

Thus the investor's risk premium associated with the uncertain return realized from the combined portfolio is

$$\pi_{rp} = \left(\frac{N_1}{W}\right)^2 \pi \tag{10.28}$$

10.9 Appendix 2: Designing Xetra

This appendix provides further detail on the development and design of Deutsche Börse's electronic trading platform, Xetra. The first steps in designing Xetra involved specifying principles that the new market should implement, and defining the system's functionality. This was done by Deutsche Börse working together with key market participants. Most importantly, the new market system was to provide equal and decentralized access to all its participants. Further, the system's functionality and the market information delivered to users (whether pre- or post-trade) were to be the same for all traders. A trader's location should not matter.

Equity trading in the German market has been and continues to be order driven. This was true both for IBIS and for floor trading that was managed by a Kursmakler acting in the capacity of auctioneer, broker, and dealer. It was clear from the beginning that Xetra should run an open limit order book (open in the sense that aggregated order volume is displayed at all price points in the order book). Additionally, order matching was automated and trader anonymity ensured.

Core features of an electronic trading system are determined by the market structure that it implements. The structure defines how orders are handled and translated into trades and transaction prices. Xetra's market model comprises diverse submodels, each with a single trading modality, or a combination of multiple modalities (i.e., it is a hybrid). Most importantly, Xetra implements both continuous trading and periodic call auction trading. This differentiation is required to cope with liquidity differences among stocks, and different liquidity needs among users depending on the size of their orders and motives for trading. The market for all stocks opens and closes with a call auction, while less liquid stocks trade in multiple call auctions per day.

Once the building blocks were defined (i.e., continuous trading and call auctions), and their combinations specified, the next design step was to detail the specific features of each of the modalities. Those features are either static (i.e., represent basic structures such as the order book) or dynamic (i.e., define processes and behavior such as order matching). The next two sections of this appendix consider the systems design in more detail for continuous trading and periodic call auction trading, respectively.

10.9.1 Continuous Trading

By the mid-1990s, order books for continuous trading with price and time priorities had been implemented around the globe. In designing Xetra, Deutsche Börse's market architects could refer to a wide range of existing examples, and to a broad microstructure literature. Once the eligible order types were identified, the center piece of the development was the definition of the detailed rules of price-time matching. The complexity of this definition was broken down into a finite set of individual cases that involved various order book situations combined with various incoming orders, for which the trading outcome was to be defined by a rule. All rules collectively described the dynamics of order matching.

A major challenge in designing continuous trading involves the measures that should be taken to provide an orderly market in periods of sharply elevated price volatility. To deal with this, the concept of a "price corridor" was formulated. Diverse corridors around historical prices were defined that set the benchmark for an "orderly" price for the next trade. If a price occurred that lay outside its corridor, trading was to be halted (briefly) with the entire order book transported into a call auction. The purpose of the call was to allow the market to consolidate in both space and time. Trading in the continuous market was resumed upon completion of the call.

Lastly, all trading parameters for the continuous platform had to be determined. This included specifying tick sizes, breadth of the price corridors, durations, and timings. Together, this provided a comprehensive overview of the "steering wheels" for the newly designed market.

10.9.2 Call Auction Trading

The purpose of Xetra's call auctions is threefold: (1) to open and close continuous trading, (2) to trade less liquid stocks in multiple calls per day with no continuous trading offered, and (3) to stabilize the market in times of large price moves. Despite those multiple purposes, a single design was defined for the auctions. Additionally, certain key consistencies between continuous trading and the call had to be achieved. For example, both limit and market orders that could be submitted to continuous trading were allowed entry into the call order book. This seemingly simple requirement was complicated to implement because it expanded the universe of possible order book configurations (and therefore necessitated more complex matching rules). Additional procedures for setting the clearing price were also required to guard against erroneous pricing that could be caused by market orders overpowering an insufficient number of limit orders. As with continuous trading, price and time priority execution rules were stipulated.

Most crucial was the degree of transparency that the calls would offer. Sufficient information about the order book had to be delivered for market participants to have relevant price and quantity information concerning actual market situations, but detailed information was suppressed to inhibit excessive information leakage and gaming. The pre-call information now available in Xetra is the highest bid and the lowest offer posted in the call when these orders do not cross, or the indicative call auction price that is calculated when the order book is crossed. In other words, the full order book content is not visible—pre-call, the Xetra screen displays only the potential outcome of the call at each point in time.

When Xetra was under development, call auction trading at prespecified times was managed on the floor by specialists who were responsible for price determination, timing, and provision of dealer liquidity. The challenge was to reengineer the call so that it could be run by a computer, not by a human intermediary. The issue that Deutsche Börse was facing was also grappled with by market microstructure academicians and other market architects. Substantial external guidance was received in the planning process. In particular, important inputs were obtained concerning the optimal degree of transparency for the call's anti-gaming measures. The availability at the time of a variety of different call auction designs (both used and proposed) enabled Xetra's calls to be designed relatively quickly.

10.9.3 Electronic Trading for Less Liquid Stocks

Kursmaklers (specialists) on the Frankfurt floor (both today and in the past) provide immediate liquidity at times when external liquidity is insufficient. The desire was strongly expressed, with two provisos, for a market maker to be incorporated into Xetra's order-driven model for less liquid stocks. The two provisos were that (1) market participants must all have equal access to information, and (2) equal access to functionality must be maintained at a maximum level. Consequently, any changes that would favor the dealers were kept to a minimum.

The dealers were referred to as "designated sponsors." Like market makers in general, the designated sponsors were given both privileges and obligations. The primary obligation is that, on request of other market participants, the designated sponsor must provide quotes for a minimum volume and maximum spread in a stock during continuous trading. Additionally, multiple designated sponsors were included, so that they might compete with each other. Concurrently, the fulfillment of each sponsor's obligation is measured, and the results are published.

The designated sponsors' primary privilege is that they can see the identity of the quote requesters in an environment that otherwise ensures complete anonymity. Further, a sponsor balances the order book in all call auctions for the stocks that it is registered in. This gives the designated sponsors a last mover advantage (the freedom to trade against any imbalance that might exist at the market clearing price). With this privilege, a designated sponsor can influence the clearing price so as to execute orders that otherwise would not have transacted in that call. Lastly, the designated sponsors, depending on their measured performance, receive fee reductions.

10.9.4 Xetra's Implementation and the Migration of Liquidity to Xetra Since 1997

Xetra went operational in Fall 1997. At the beginning, the new system attracted roughly 60 % of trading in the most liquid segment of the market, the 30 DAX stocks. Trading on Xetra for mid-cap stocks was not as successful—market share for this segment of the market was about 20 %, as the less liquid stocks largely continued at that time to trade on the floor. But the 1997 launch was just the start of a sequence of releases that have continued through the current time.

One more recent innovation was the "continuous call auction." With this facility, calls are not held at prespecified times but are triggered by the occurrence of a "critical" liquidity situation. The continuous call comprises a dealer-auctioneer who is responsible for providing a base level of liquidity in each call, as well as controlling its timing. Additionally, Xetra allows internalization of trading by member firms. Consequently, Xetra, which originally started as an exchange trading system, now also serves as the technical platform for OTC trading.

Major innovations have benefited a broad range of cap sizes and, across the board, floor trading has continued to decline. Xetra has now been rolled out to 260 member firms in Europe, and its market share currently stands at 95 % of all on-exchange trading in Germany today.

References

1. Francioni, Reto, Sonali Hazarika, Martin Reck and Robert A. Schwartz. "Equity Market Microstructure: Taking Stock of What We Know." *Journal of Portfolio Management*, Fall, 2008, forthcoming.
2. Schwartz, Robert and Reto Francioni. *Equity Markets in Action*, John Wiley and Sons, 2004.
3. Bernstein, Peter. "The Surprising Bond between CAPM and the Meaning of Liquidity." *Journal of Portfolio Management,* Fall 2007.
4. Ho, Thomas, Robert Schwartz, and D. Whitcomb. "The Trading Decision and Market Clearing Under Transaction Price Uncertainty." *Journal of Finance,* March 1985.
5. Cohen, Kalman, Steven Maier, Robert Schwartz, and David Whitcomb. "Market Makers and the Market Spread: A Review of Recent Literature." *The Journal of Financial and Quantitative Analysis*, 1979, 14, pp. 813–835.
6. O'Hara, Maureen. Market Microstructure Theory. Basil Blackwell, 1997, Cambridge, MA.
7. Madhavan, Ananth. Market Microstructure. *Journal of Financial Markets*, 2000, 3, pp. 205–258.
8. Biais, Bruno, Bruno, Larry Glosten, and Chester Spatt. "Market microstructure: A Survey of Microfoundations, Empirical Results, and Policy Implications." *Journal of Financial Markets*, 2005, 8, pp. 217–264.
9. Parlour, Christine and Duane Seppi. "Limit Order Markets: A Survey." Forthcoming in Handbook of Financial Intermediation & Banking edited by A.W.A. Boot and A. V. Thakor, 2008.
10. Hasbrouck, Joel. Empirical Market Microstructure. Oxford University Press, 2007.
11. Garman, Mark. "Market Microstructure." *Journal of Financial Economics*, June 1976, pp. 33–53.

12. Stigler, George. "Public Regulation of the Securities Markets." *Journal of Business*, April 1964, pp. 117–142.
13. Smidt, Seymour. "Which Road to an Efficient Stock Market: Free Competition or Regulated Monopoly?" *Financial Analyst Journal,* 27, 1971, pp. 18–20, 64–69.
14. Tinic, Sneha. "The Economics of Liquidity Services." *The Quarterly Journal of Economics*, 86, 1, 1972, pp. 79–93.
15. Tinic, Sneha, and Richard West. "The Securities Industry under Negotiated Brokerage Commissions: Changes in the Structure and Performance of New York Stock Exchange Member Firms" Bell Journal of Economics, 11, (Spring 1980), pp. 29–41.
16. Bagehot, Walter (pseudonym). "The Only Game in Town." *Financial Analysts Journal,* (Mar/April 1971), pp. 12–14, 22.
17. Stoll, Hans. "The Supply of Dealer Services in Securities Markets." *Journal of Finance*, 33, 1978, pp. 1133–1151.
18. Amihud, Yakov, and Haim Mendelson. "Dealership market: Market-Making with Inventory." *Journal of Financial Economics*, 8, 1980, pp. 31–53.
19. Ho, Thomas, and Hans Stoll. "On Dealer Markets under Competition." *Journal of Finance*, 35, 1980, pp. 259–267.
20. Ho, Thomas, and Hans Stoll. "Optimal Dealer Pricing under Transactions and Return Uncertainty." *Journal of Financial Economics*, 9, 1981, pp. 47–73.
21. Ho, Thomas, and Hans Stoll. "The Dynamics of Dealer Markets Under Competition." *Journal of Finance*, 38, 1983, pp. 1053–1074.
22. Mildenstein, Eckart, and Harold Schleef. "The Optimal Pricing Policy of a Monopolistic Marketmaker in the Equity Market." *Journal of Finance*, 38, 1983, pp. 218–231.
23. Stoll, Hans. "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests," *Journal of Finance*, 44, 1989, pp. 115–134.
24. Demsetz, Harold. "The Cost of Transacting." *The Quarterly Journal of Economics*, 82, 1, 1968, pp. 33–53.
25. Glosten, Lawrence, and Paul Milgrom. "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents." *Journal of Financial Economics*, 14, 1985, pp. 71–100.
26. Kyle, Albert. "Continuous Auctions and Insider Trading." *Econometrica*, 53, 1985, pp. 1315–1335.
27. Grossman, Sandford, and Joseph Stiglitz. On the Impossibility of Informationally Efficient Markets. *American Economic Review*, 70, 3, 1980, pp. 393–408.
28. Milgrom, Paul and Nancy Stokey. "Information, Trade and Common Knowledge." *Journal of Economic Theory*, 26, 1, 1982, pp. 17–27.
29. Copeland, Thomas, and Dan Galai. "Information Effects on the Bid-Ask Spread." *Journal of Finance*, 38, 1983, pp. 1457–1469.
30. Easley, David, and Maureen O'Hara. "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics*, 19, 1987, pp. 69–90.
31. Easley, David, and Maureen O'Hara. "Order Form and Information in Securities Markets." *Journal of Finance* 46, 1991, pp. 905–928.
32. Easley, David, and Maureen O'Hara. "Time and the Process of Security Price Adjustment." *Journal of Finance*, 47, 1992, 577–606.
33. Mendelson, Morris, Junius Peake and T. Williams, "Toward a Modern Exchange: The Peake-Mendelson-Williams Proposal for an Electronically Assisted Auction Market," in E. Bloch and R. Schwartz, eds. *Impending Changes for Securities Markets: What Role for the Exchange?* JAI Press, 1979.
34. Domowitz, Ian and Benn Steil. "Automation, Trading Costs, and the Structure of the Securities Trading Industry." Brookings-Wharton Papers on Financial Services, 1999, 33–92.
35. Cohen, Kalman, Steven Maier, Robert Schwartz, and David Whitcomb. "Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread." *The Journal of Political Economy*, April 1981, pp. 287–305.

36. Glosten, Lawrence. "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49, 1994, pp. 1127–1161.
37. Handa, Puneet, and Robert Schwartz. "Limit Order Trading." *Journal of Finance*, December 1996a, pp. 1835–1861.
38. Foucault, Thierry. "Order Flow Composition and Trading Costs in a Dynamic Order Driven Market." *Journal of Financial Markets,* 2, 1999, pp. 99–134.
39. Parlour, Christine. "Price Dynamics in Limit Order Markets." *Review of Financial Studies,* 11, 1998, pp. 789–816.
40. Handa, Puneet, Robert Schwartz, and Ashish Tiwari. "Quote Setting and Price Formation in an Order Driven Market." *Journal of Financial Markets*, 2003, 6, pp 461–489.
41. Foucault, Thierry, Ohad Kaden and Eugene Kandel. "The Limit Order Book as a Market for Liquidity." *Review of Financial Studies,* 18, 2005, pp. 1171–1217.
42. Goettler, Ronald, Christine Parlour, and Uday Rajan. "Equilibrium in a Dynamic Limit Order Market." *Journal of Finance,* 60, 2005, pp. 2149–2192.
43. Grossman, Sandford. "The Information Role of Upstairs and Downstairs Markets." *Journal of Business*, 1992, 65, pp. 509–529.
44. Keim, Donald, and Ananth Madhavan. "The Upstairs Market for Large-Block Transactions: Analysis and Measurement of Price Effects." *Review of Financial Studies,* 9, 1996, pp. 1–36.
45. Madhavan, Ananth, and Minder Cheng. "In Search of Liquidity: An Analysis of Upstairs and Downstairs Trades." *Review of Financial Studies*, 10, 1997, pp. 175–204.
46. Seppi, Duane. "Equilibrium Block Trading and Asymmetric Information." *Journal of Finance*, 45, 1990, pp. 73–94.
47. Porter, David and Daniel Weaver. "Post-Trade Transparency on Nasdaq's National Market System." *Journal of Financial Economics*, 50, 2, 1998, pp. 231–252.
48. Engle, Robert, and Clive Granger. "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica*, 55, 1987, pp. 251–276.
49. Angel, James. "Tick Size, Share Prices, and Stock Splits." *Journal of Finance*, 52, 1997, pp. 655–681.
50. Harris, Larry. "Stock Price Clustering and Discreteness." *Review of Financial Studies*, 4, 1991, pp. 389–415.
51. Harris, Larry. "Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes." *Review of Financial Studies*, 7, 1994, pp. 149–178.
52. Domowitz, Ian, Jack Glen and Ananth Madhavan. "Liquidity, Volatility, and Equity Trading Costs across Countries and Over Time." *International Finance* 4, 221–256, 2001.
53. Hasbrouck, Joel. "Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement." *Review of Financial Studies,*, 6 (1), 1993, pp. 191–212.
54. Hasbrouck, Joel, and George Sofianos. "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists." *Journal of Finance*, 48, 1993, pp. 1565–1593.
55. Menkveld, Albert, Siem Koopman and André Lucas. "Modelling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Methods." *Journal of Business and Economic Statistics,* 25 (2), 2007, pp. 213–225.
56. Paroush, Jacob, Robert Schwartz, and Avner Wolf. "The Dynamic Process of Price Discovery in an Equity Market." Working paper, Baruch College, CUNY, 2008.
57. Barclay, Michael and Terrence Hendershott and Timothy McCormick. "Competition Among Trading Venues: Information and Trading on Electronic Communications Networks." *Journal of Finance*, 58, 2003, pp. 2637–2666.
58. Bessembinder, Hendrik, and Herbert Kaufman. "A Comparison of Trade Execution Costs for NYSE and Nasdaq-listed Stocks." *The Journal of Financial and Quantitative Analysis*, 32, 1997a, pp. 287–310.
59. Bessembinder, Hendrik, and Herbert Kaufman. "A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks." *Journal of Financial Economics*, 46, 1997b, pp. 293–319.
60. Bessembinder, Hendrik. "Trade Execution Costs on Nasdaq and the NYSE: A Post-Reform Comparison." *The Journal of Financial and Quantitative Analysis*, 34, 1999, pp. 387–407.

61. Bessembinder, Hendrik. “Quote-Based Competition and Trade Execution Costs in NYSE-Listed Stocks”, *Journal of Financial Economics*, 70, 2003, pp. 385–422.
62. Christie, William, Jeffrey Harris, and Paul Schultz. Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes? *Journal of Finance*, 49, 1994, pp. 1841–1860.
63. Christie, William, and Paul Schultz. Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes? *Journal of Finance*, 49, 1994, pp. 1813–1840.
64. Hasbrouck, Joel. “Measuring the Information Content of Stock Trades.” *Journal of Finance*, 46, 1991, pp. 179–207.
65. Hasbrouck, Joel. “One security, many markets: Determining the Contribution to Price Discovery.” *Journal of Finance*, 50, 1995, pp. 1175–1199.
66. Biais, Bruno., Pierre Hillion, and Chester Spatt. “An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse.” *Journal of Finance*, 50, 1995, pp. 1655–1689.
67. Ozenbas, Deniz, Robert Schwartz, and Robert Wood. “Volatility in U.S. and European Equity Markets: An Assessment of Market Quality,” *International Finance,* Volume 5 Number 3, Winter 2002, pp. 437–461.
68. Sandas, Patrick. “Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market.” *Review of Financial Studies*, 14, 2001, pp. 705–734.
69. Duffie, Darrell, Lasse Pedersen and Nicolae Garleanu. “Valuation in Over-the-Counter Markets.” *Review of Financial Studies*, 20, 2007, pp. 1865–1900.
70. Bernstein, Peter. “Liquidity, Stock Markets and Market Makers.” *Financial Management,* Summer 1987, pp. 54–62.
71. Miller, Edward, “Risk, Uncertainty and Divergence of Opinion,” *Journal of Finance,* 1977, 32, 1151–1168.
72. Davis, Paul, Michael Pagano, and Robert Schwartz. “Divergent Expectation.” *Journal of Portfolio Management*, Fall 2007, forthcoming.
73. Davis, Paul, Michael Pagano, and Robert Schwartz. “Life After the Big Board Goes Electronic.” *Financial Analysts Journal*, Volume 62, Number 5, September/October 2006, pp. 14–20.
74. Beja, Avraham and Nils H. Hakansson. “Dynamic Market Processes and the Rewards to Up-to-Date Information.” *Journal of Finance*, 32, 1977, pp. 291–304.
75. Beja, Avraham and M. Barry Goldman. “On the Dynamic Behavior of Prices in Disequilibrium.” *Journal of Finance*, 35, 1980, pp. 235–248.
76. Copeland, Thomas. “A Model of Asset Trading under the Assumption of Sequential Information Arrival.” *Journal of Finance*, 31, 1976, pp. 1149–1168.
77. Back, Kerry and Shmuel Baruch. “Working Orders in Limit-Order Markets and Floor Exchanges.” *Journal of Finance*, 62, 2007, pp. 1589–1621.
78. Economides, Nicholas, and Robert Schwartz. “Electronic Call Market Trading.” *Journal of Portfolio Management*, Spring 1995, pp. 10–18.
79. Pagano, Marco and Ailsa Röell. “Transparency and Liquidity: a Comparison of Auction and Dealer Markets with Informed Trading”, *Journal of Finance*, 51, 1996, pp. 579–611.
80. Harvey, Campbell and Roger Huang. “Volatility in the Foreign Currency Futures Market.” *Review of Financial Studies,* 4, 1991, pp. 543–569.
81. Jones, Charles, Gautam Kaul and Marc Lipson. “Information, Trading and Volatility.” *Journal of Financial Economics*, 36, 1994, pp. 127–154.
82. Bessler, Wolfgang. Editor, Bösen, Banken und Kapitalmärkte, Duncker & Humblot, 2006.
83. Amihud, Yakov and Haim Mendelson. “Asset Pricing and the Bid-Ask Spread.” *Journal of Financial Economics*, 17, 1986, pp. 223–249.
84. Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam. “Commonality in Liquidity.” *Journal of Financial Economics*, 56, 2000, pp. 3–28.
85. Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam. “Liquidity and Market Efficiency.” *Journal of Financial Economics*, 87, 2008, pp. 249–268.
86. Hasbrouck, Joel and Duane Seppi. “Common Factors in Prices, Order Flows and Liquidity.” *Journal of Financial Economics*, 59, 2001, pp. 383–411.

87. Amihud, Yakov. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets*, 5, 2002, pp. 31–56.
88. Pástor, Luboš and Robert Stambaugh. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 113, 2003, pp. 642–685.
89. Harris, Larry. Trading and Exchanges: Market Microstructure for Practitioners. 2003, Oxford University Press.
90. Handa, Puneet and Robert Schwartz. "How Best to Supply Liquidity to a Securities Market." *Journal of Portfolio Management*, Winter 1996b, pp. 44–51.
91. Schwartz, Robert. *Reshaping the Equity Markets: A Guide for the 1990s,* Harper Business, 1991.

Chapter 11
Exchanges: Link to the Real Economy

Michael Heise

11.1 The Stock Markets Back in the Day: A Source and Instrument of State Power

As reflected in a book by the British economic historian Niall Ferguson [1], the story behind the ascent and decline of nations can also be told as a story of their stock exchanges and financial markets. As competing nations went head to head in the battle for power, the ability to finance wars and buy influence proved to be hugely important: access to capital markets and financial innovations were a key strategic advantage for kings and rulers alike.

The Italian city-states, for example, laid the foundation stone for their success and for the financial system as we know it today when they introduced state-of-the-art bookkeeping methods. The invention of the "public company" allowed the Netherlands and the UK to raise the vast amounts required to finance their trade empires. The development of America's Wild West would have been impossible had it not been for the establishment of specialized investment banks, because the funds required to build the country's transcontinental railroad routes could presumably never have been raised using conventional methods.

But the financial endeavors were not always success stories: both France and Scotland tried to use speculation on trade with overseas lands in order to realize their colonial ambitions and solve their budget problems.[1] Both kingdoms, how-

[1] In 1695 the Scottish Parliament founded the Company of Scotland Trading to Africa and the Indies and Scots invested £400,000 in the new-founded company which aimed to establish a colony in the area of today's Panama. About 1200 people set off to Panama in 1698 but only a quarter returned. The failed adventure wiped out capital equal to half the gross domestic product of Scotland at that time. The French colonial ambition of that day culminated in the Mississippi Bubble which will be discussed in Sect. 11.4.

M. Heise (✉)
Allianz SE, Königinstraße 28, Munich 80802, Germany
e-mail: michael.heise@allianz.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_11

ever, failed spectacularly, leaving a lasting mark on their strategic position in the competition between Europe's powers. Scotland's thwarted ambitions are actually likely to have ended Scottish independence.

So the stock markets have a long and eventful history behind them. Shares were traded publicly as long ago as in the era of ancient Rome. The key event in modern stock exchange history, however, is the establishment of the Dutch East India Company (*Vereenigde Oostindische Compagnie*, or *VOC* for short) in 1602. Associations of traders specializing in trade with Southeast Asia had sprung up to cover the duration of a trading voyage. These expeditions lasted around 14 months on average and were extremely risky: in some cases, only half of the ships ever made it back to Europe. In this sort of situation, bringing various parties together to bear the capital and risk involved was a logical step.

The Dutch East India Company was not, however, designed purely to cover the duration of one expedition, but aimed to challenge the supremacy of the trading posts belonging to the rival maritime powers Portugal and Spain. As a result, it was set up for a period of 21 years from the outset and was granted a monopoly on trading with Southeast Asia. With initial capital of 6.45 million Dutch guilders, the Dutch East India Company was the biggest company of its time—around eight times as big as its English rival, the East India Company. The structure of the company meant that investors were only entitled to claim back their invested capital after a period of 10 years. Investors who needed their money back sooner had no option but to sell their shares. This laid the foundation for the modern-day stock markets. The Dutch East India Company was so successful that, in the period leading up to 1650, shareholders received annual dividends in excess of 16 %. The Amsterdam Stock Exchange was set up only just after the VOC to provide a platform for the flourishing trade in the company's shares. This move was followed, in 1609, by the establishment of the first central bank of the modern era, which naturally accepted VOC shares as loan collateral.

11.2 The Stock Markets Today: A Catalyst to Economic Advancement

Over the course of time, the stock exchange proved to be a resounding success: today, stock exchanges are the backbone of any developed economy. The prospect of a modern economy without highly developed stock exchanges now seems inconceivable. Although this success story started in Amsterdam, it was ultimately the London Stock Exchange that emerged as the global leader, boosted by the British Empire. For a long time, the many regional European stock markets merely served as a stepping stone to a listing on the London Stock Exchange. In the twentieth century, the focus then shifted even further west: today, the two major US stock exchanges, the New York Stock Exchange and the Nasdaq, are bigger, in terms of the market capitalization of the companies listed there, than the leading stock markets of Europe and Asia combined.

A few figures from the recent past highlight just how important the role played by the stock markets has become. The deregulation policy that followed the period of stagflation in the 1970s enhanced the economic importance of the capital markets: in the period between 1988 and 2007, the market value of all publicly traded companies rose almost sixfold from roughly USD 12 trillion to more than USD 64 trillion. During the same period, global economic output approximately just tripled. So even if we account for rising incomes, the importance of exchanges as a source of funding has greatly increased: the ratio of corporate market capitalization to global economic output almost doubled, as it climbed from 66 to 111 %. Although the financial crisis slammed the brakes on the stock exchange surge, it was unable to stop it entirely. It only took until 2013 for global market capitalization to bounce back to its precrisis level, although it has not yet returned to its previous highs in relation to global GDP (see Fig. 11.1).

Although the US stock markets are the unchallenged leaders in absolute terms, a look at the relationship between stock markets and economic output paints a different picture: surprisingly, this sort of analysis puts South Africa at the top of the table, followed by Singapore and Switzerland in second and third places, respectively; the USA follows in fifth place, with Germany coming in at only 29th. China finds itself only on place 18. But this relative ranking belies the sheer size of its stock market: China (incl. Hong Kong) has the second biggest stock market behind the USA in absolute numbers with a total market capitalization of USD 9.2 trillion. Nonetheless, up to now, China's financial system has been rather bank dominated.

Generally, one thing that is striking is that the top places on the ranking list are not populated exclusively by developed economies. Rather, half of the countries in the top ten are emerging markets. This is an impressive evidence of the global progress made by the stock markets and the role that they play in economic advancement (see Fig. 11.2).

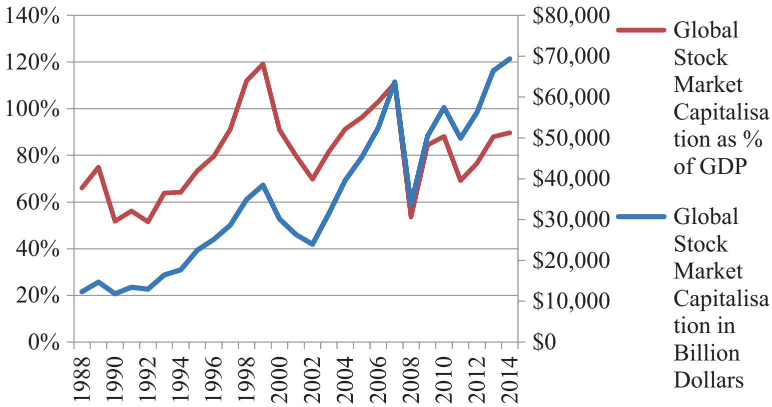


Fig. 11.1 Global stock market capitalization 1988–2014 (source: World Federation of Exchanges, World Economic Outlook Database (April 2015))

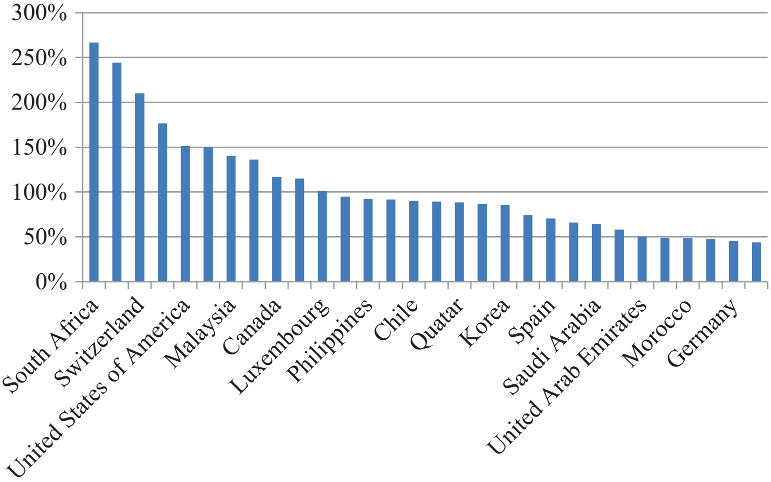


Fig. 11.2 Domestic stock market capitalization as % of GDP in 2014 (note: Scandinavia and Baltics include Denmark, Finland, Sweden, Iceland, Estonia, Latvia, and Lithuania (OMX Exchanges); Western Europe includes Belgium, France, the Netherlands, and Portugal (Euronext). Source: World Federation of Exchanges, World Economic Outlook Database (April 2015))

If we look at the number of companies traded, the ranking list of the world's largest stock exchange changes again. This list sees India lead the field, with almost 5200 publicly traded companies, around 1000 more than in the USA. With 665 companies, Germany nevertheless manages to secure 13th place, behind the much smaller economies of Canada, Spain, Serbia, and Malaysia.

All in all, however, Germany has certainly earned the title of stock market cynic: the size of the country's stock market does not adequately reflect its economic power. One root cause can certainly be found in the low proportion of shareholders in Germany: only around 13 % of the German population holds direct or indirect equity investments. This is significantly less than in the USA (56 %) or Japan (27.7 %) for example; even the country's neighbor Switzerland has a slightly higher proportion of shareholders, at 19.4 %.[2]

This is an unsatisfactory development both from the view of savers wealth accumulation and from the view of the corporate sector. As said in the introduction, the main reason why stock exchanges were set up in the first place was the need to draw upon broader sections of the population to finance the corporate sector, e.g., costly trading voyages or inventions like the railway. The huge appetite for capital had an eminently positive side effect: stock exchanges gave a larger number of citizens access to the financial markets, enabling broad sections of the population to participate in the economy's productive assets.

This particular stock market role to allow long-term asset accumulation by broader sections of the population has become even more important today, in a

[2] All figures quoted from Deutsches Aktieninstitut (DAI, 2015).

world dominated by aging societies. At the same time, growing participation in the capital markets also makes a significant contribution to the success of an economy as a whole. Broad and deep capital markets are essential for economic growth, because they put the population's savings to productive use.

11.3 The Connection Between Exchanges and the Real Economy

Financial markets perform several key tasks in this capital allocation process: first and foremost they provide a mechanism by which savers (primarily private households) can lend their savings to borrowers (primarily governments and firms) to invest in productive projects. Financial markets also allow a transfer of risk between participants with different risk-bearing capacities and permit savers to diversify their asset portfolios, for example by using exchange-traded instruments like catastrophe bonds or credit default swaps. And by keeping the costs of transactions as low as possible, these markets promote economic efficiency.

Functioning capital markets would give medium-sized and young growth companies financing options over and above the conventional routes. Whereas in the past it was common practice to finance a start-up using one's savings or with help from family members, as access to bank loans once was granted only when the business was up and running successfully, modern instruments such as venture capital allow the conventional start-up process to be turned on its head: even without a stable cash flow or loan collateral, start-ups can also seek financing via the capital markets (Fig. 11.3).

The key factor in this process is not so much the volume of capital moved by the financial markets, but rather their ability to channel funds to those parties that can put them to the best use. Identifying those borrowers hinges on the ability of financial markets to accurately pool and provide information. In the best case they provide the best available estimate of a firm's risk and reward perspectives by processing new information, mediating between the different expectations of market participants, and, in doing so, determining market clearing prices. The most important characteristic of financial markets therefore is the quality of their allocation [2]. In this respect, capital markets and in particular stock markets seem to have the edge over banks, at least in more advanced markets: A recent study by the OECD has shown that in most OECD countries an increase in intermediated credit might lead to even *slower* growth whereas stock market expansion normally stimulates growth ([3], p. 17).

On the other hand, underdeveloped countries often are at a general disadvantage because their financial systems do not perform these tasks, information processing and risk transfer, efficiently. Instead, they often channel capital into unproductive, but well-connected, sectors. Malfunctioning capital markets primarily benefit privileged citizens and do not give the broad population access to productive capital. There is a well-documented relationship linking financial development, or more precisely the quality of capital allocation, to the economic development of a country. Furthermore,

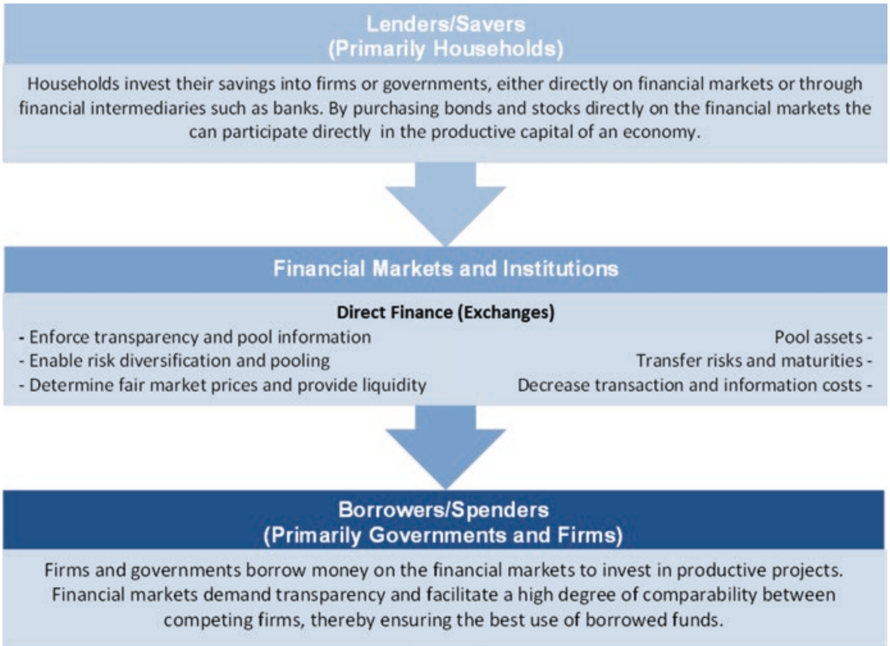


Fig. 11.3 Flow of funds through the financial markets (Source: Own illustration)

firm-level evidence shows that better developed financial systems ease external financing constraints facing firms [4]. Based on those empirical observations it is quite safe to say that efficient capital markets and access to the international markets are a vital condition for the economic development for the world's poorer countries.[3]

Contrary to what one might believe in the aftermath of the financial crisis, functioning capital markets can also help to stabilize the economy. Especially in times of crisis, capital markets are a key source of corporate financing when banks are no longer available as lenders. Between 2009 and 2013, for example, the drop in the number of corporate loans in the Eurozone was matched almost exactly by an increase in new corporate bond issues [5].

11.4 Rational Expectations and Irrational Excesses

The triumph of the capital markets has its downsides as well. Only one century after the establishment of the Dutch East India Company, the Scot John Law, who was based in France, triggered what is likely to be the very first speculative stock market

[3] However, for developed markets the relationship between financial development and growth seems to be more nuanced. Beyond a certain threshold, even more credit by banks can be detrimental to economic growth [3].

bubble in history. After setting up a private central bank with banknote privileges with the support of the French Regent in 1715, he set up the Mississippi Company, which was modeled on the Dutch East India Company and had a monopoly on trading with the French colonies of North America. The required capital originally came from devalued government bonds, which had previously been absorbed by John Law's central bank in de facto terms. In order to create demand for shares in the company, the central bank, which was controlled by Law, granted loans and accepted the shares as collateral. In today's world, this would, according to Ferguson [1], be tantamount to a scenario in which one person controlled the US Fed and, at the same time, all of the companies listed on the S&P500.

Lured by the promise of the treasures tucked away in the French colonies, capital started to flow to France from abroad as well. At the height of the bubble, shares in the Mississippi Company had gained 2000 %. When the promises turned out to be empty, however, many investors started to pull their money out, effectively bursting the bubble. The massive asset losses that ensued plunged France into a deep crisis. The irony of this story is that many investors fled and "invested" their money in the British South Sea Company—creating the next bubble right away. The British bubble then burst only a few months after the French one, although the consequences were much less drastic.

And these were anything but the last bubbles, as Charles Kindleberger showed in his [6] book entitled "Manias, Panics, and Crashes." All speculative bubbles appear to follow the same script: first, there is a change in the overall economic framework that allows new business models to emerge, be it due to the discovery of new reserves of natural resources, new technologies, or regulatory changes. The higher profit expectations are soon reflected in rising stock market prices, which sooner or later overshoot the level that can be deemed appropriate, attracting a large number of small investors. Slowly but surely, however, it becomes clear that the profit expectations will not materialize, and the first lot of insiders start to pull their money out. As soon as the share prices start to drop, the vast majority of investors sell their shares as well, bursting the bubble completely. In most cases, stock market crises also have a direct impact on the real economy.

But is the term "bubble" really appropriate to describe regularly recurring stock exchange highs and crashes? Not according to Eugene Fama, one of the winners of the 2013 Nobel Prize in economics. As long as there is sufficient liquidity available on the market and no misdirected incentives from the regulatory side or anywhere else, market participants price new information into stock market prices as soon as it becomes available; in such cases, the markets are "information efficient." Pricing on the stock markets and other stock exchanges can be described as a process of discovery: each new piece of information that becomes available and has an impact on the return prospects or risk associated with a particular investment is exploited on the markets as soon as possible. Ultimately, the price of an investment represents the best possible assessment of its opportunities and risks. This is why, according to Fama, bubbles should never emerge in the first place. What is generally described as a bubble, for example the hi-tech bubble on the *Neuer Markt* in Germany, could also be interpreted as a rational reaction to the uncertain profit outlook associated with new technologies [7].

11.5 Asset Prices and Monetary Policy

Although there is no doubt that Fama's line of argumentation is convincing from a theoretical standpoint, the sort of efficiency he postulates is something we are unlikely to encounter in reality. This is because "artificial" disincentives standing in the way of efficient information processing are, sadly, an all-too-common occurrence. One particular "misdirected incentive" has been thrust into the spotlight of late: monetary policy and zero interest rates.

As excessive developments on the stock market tend to coincide with periods of relatively loose monetary policy, it could be central banks that promote mispricing with excessively loose monetary policy. This is the view taken by researchers at the Bank for International Settlements (BIS). Borio et al. [8] argue that the risk of bubbles increases when monetary policy is too loose, because it encourages financial market players to take more risks. After all, many (institutional) investors are bound by certain yield requirements and start taking bigger risks when yields on fixed-income investments head south. At the same time, however, they are better placed to acquire risks during an upswing as well, because low interest rates boost the value of their loan collateral. These patterns, which also have a regulatory background, promote the pro-cyclical nature of the capital markets and make markets fall all the faster when a downturn hits.

But boom-and-bust cycles are not just a problem that affects the stock markets. Their negative impact is likely to radiate into other areas, too. Economists from the BIS, for example, point out that productivity in other sectors of the economy starts to fall as soon as the financial sector becomes too big [9]. One reason for this could be because the financial sector poaches the most talented employees from other sectors. Another is the preference among financial institutions for companies that can furnish loan collateral. But the assets of innovative companies that spend a lot of money on research and development are largely intangible, meaning that they are not suitable for use as loan collateral. This could put them at a disadvantage compared to less productive companies, such as those in the real estate sector. Although the latter can provide large loan collateral, they do not contribute strongly to growth in overall productivity.

So it is no wonder that, since the financial crisis, the focus has been less on the positive macroeconomic effects of stock exchanges and capital markets, and more on the question as to how excessive ups and downs can be contained. Financial market regulation is one of the only issues. Furthermore it is debated what role the prices of shares, bonds, real estate, and other assets should play in central bank decisions. As long as asset prices provide an accurate reflection of all of the available information on fundamental factors such as inflation and growth, which the central banks take into account anyway, there is no specific role for asset prices. But what if bubbles can emerge on the asset markets, should the central bank not then try to stand in their way?

In the days prior to the financial crisis, many economists believed that a study conducted by Bernanke and Gertler [10] provided the answer to this question. Central banks should not try to combat irrational exaggerated developments on the financial markets. Instead, they should merely ensure that the economy does not slip

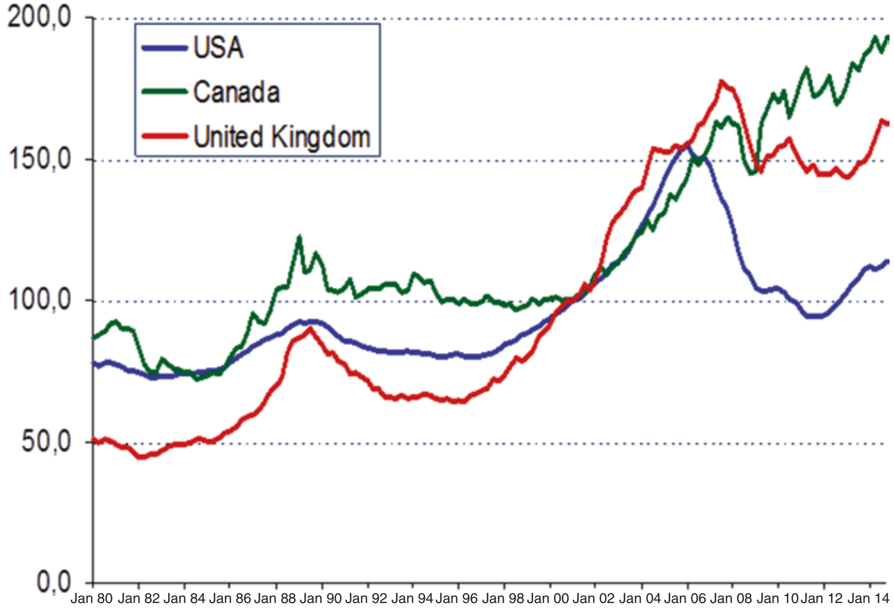


Fig. 11.4 Real residential property prices for the USA, the UK, and Canada (Q1 2001 = 100) (Source: FRED Economic Data)

into a recession as soon as a bubble bursts. Bernanke and Gertler argue that it is extremely difficult to identify a bubble in time at all. Whereas in hindsight the US property bubble can easily be recognized as a bubble, property prices in Canada and the UK also charted a similar rise in the period between 1990 and 2007. But the USA was the only country that witnessed such a sharp correction (see Fig. 11.4).

But even if it were to detect excessive valuations early on, it is not clear whether a central bank could actually do anything about it. Excessive valuations like those seen on Germany's *Neuer Markt* are normally characterized by such exaggerated profit expectations that upping key interest rates by a few percentage points would hardly have any impact. Real estate prices depend primarily on long-term interest rates—which, however, can only be influenced indirectly by (conventional) monetary policy.

Since the financial crisis, however, this doctrine has been challenged. Strongly rising asset prices tend to exaggerate financial imbalances. Research conducted by Jordà et al. [11] points to credit cycles, and not asset prices, as the real problem. In boom times, when collateral is abundant, banks and shadow banks tend to grant excessive loans that pose a threat to the stability of the financial system and the real economy. From this angle, rising asset prices are merely a symptom of a more fundamental problem. The transmission of monetary policy through the credit channel and the portfolio channel can be mutually reinforcing. Therefore monetary policy actually should react to financial market booms by leaning against the wind of excessive asset price developments. Hereby monetary policy should be supported by targeted macroprudential measures.

Regulatory intervention has also been employed to restrict the capital markets' volatility. But often it has unintended consequences. There is, for example, evidence to suggest that the ban on short selling imposed during the financial crisis significantly reduced the liquidity in trading with small cap or high-risk shares, posing an obstacle to pricing in the process [12]. Another example is the ban on proprietary trading by banks. The rule that has come to be known as the Volcker Rule in the USA means that, in some markets, banks can no longer act as market makers. This reduces the liquidity on these markets and can trigger major price fluctuations. In other words: the efforts to restore stability can end up being a direct path to higher volatility.

11.6 Conclusion

Over the course of the centuries, there has been the same fundamental role played by the stock markets. They are the way to raise risk capital and fund the growth of companies. Therefore, stock markets are of major importance today, not least in emerging markets. Also the "democratic" aspect of the stock market, namely allowing broad sections of the population to participate in an economy's productive assets, remains as important as ever. After all, long-term asset accumulation is a must for each and every one of us in aging societies. And, in a world of burgeoning debt, what could be better suited for asset accumulation than real assets? As long as the period of low interest rates continues and central banks across the globe keep doing everything to lower the cost of debt and reflate economies, there is virtually no alternative to shares that offer comparatively higher protection against inflation.

Of course, stock markets also pose some problems. The theory of efficient financial markets failed the reality check of the financial crisis. In many cases, pricing is not as information efficient as it should be—be it because of misdirected incentives by monetary policy or by financial market regulation or by exuberant financial behavior of market participants like herding. The big swings of the stock markets in the last two decades have discouraged many savers to invest in stocks, especially in lower income segments of our societies. Therefore a rising stock market overproportionately benefits the wealthier parts of society. This needs to be overcome, and stock market investments need to be broadened.

That requires a sound policy framework in terms of regulation and monetary policy. Following the immediate shock of the financial crisis and the ensuing monetary policy and regulatory reactions, it is now time to switch back from crisis mode to a long-term normalization pattern.

References

1. Ferguson, N. (2008). The ascent of money: A financial history of the world. New York, NY: Penguin.
2. Levine, Ross. "Finance, long-run growth, and economic opportunity." VOX (October 25, 2011). Accessed on May 28, 2015.

3. OECD (2015), Finance and inclusive growth, OECD Economic Policy Paper No.14, OECD 2015.
4. Levine, Ross. "Finance and growth: theory and evidence." Handbook of economic growth 1 (2005): 865–934.
5. Deutsche Bank Research. Corporate funding in Europe: Bonds replacing loans? Accessed on May 28, 2015.
6. Kindleberger, C. P. M. (1978). Panics, and Crashes: A History of Financial Crises. New York, NY: Palgrave Macmillan.
7. Fama, Eugene F. "Interview with Eugene Fama." The Region (2007): 15–22.
8. Borio, Claudio, and Haibin Zhu. "Capital regulation, risk-taking and monetary policy: a missing link in the transmission mechanism?." Journal of Financial Stability 8.4 (2012): 236–251.
9. Cecchetti, S. G., & Kharroubi, E. (2013). Why does financial sector growth crowd out real economic growth?. Bank for International Settlements Working Papers No 490.
10. Bernanke, B., & Gertler, M. (1999). Monetary Policy and Asset Price Volatility. In New Challenges for Monetary Policy. Kansas City: Federal Reserve Bank of Kansas City.
11. Jordà, Ò., Schularick, M., & Taylor, A. M. (2011). Financial crises, credit booms, and external imbalances: 140 years of lessons. IMF Economic Review, 59(2), 340–378.
12. Beber, A. & Pagano, M (2010). Short-selling bans in the crisis: A misguided policy. In Vox. Accessed on May 28, 2015.

Chapter 12
The Role of High-Frequency Trading in Modern Financial Markets

Wolfgang Eholzer and Randolf Roth

In the past decades, financial markets have undergone a profound change driven by a combination of technological advancement and fierce competition between both market participants and marketplaces. The technological advancement has moved market venues from floor to electronic venues and in that process generated substantial benefits to the investing public as cost of market entry has been drastically reduced with respect to:

- Provision of low-cost Internet-based direct electronic market access versus traditional telephone-based high-touch market access.
- Ease of access to information and therefore transparency on traded prices, available bids and offers as well as fundamentals of financial instruments have increased substantially through Internet and other electronic means.
- Markets have become more liquid in terms of displaying tight bid-offer spreads across the globe reducing the implicit transaction cost.
- Competition between trading venues and central infrastructures has reduced their fees significantly.

The largest part of this impact has materialized in the last decade, i.e., from 2000 onwards when most of the traditional floor-based markets have moved electronic and new electronic market venues such as MTFs and ECNs developed. Since the financial crisis HFT is seen by an increasing part of the public but also political and regulatory stakeholders as a source of trouble. At a first glance this seems unjustified as the reasons for the financial crisis in the USA as well in Europe had nothing to do with HFT and not a single tax Euro or USD has spent on those firms.

The purpose of this chapter is to have a more in-depth look at the role of HFT in modern electronic markets. This includes a discussion of the concerns and reservations

W. Eholzer (✉) R. Roth
Eurex Frankfurt AG, Mergenthalerallee 61, 65760 Eschborn, Germany
e-mail: wolfgang.eholzer@deutsche-boerse.com; randolf.roth@eurexchange.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_12

towards HFT. The chapter is structured as follows: Part 1 discusses the definition of HFT and differentiates it from algorithmic trading in general. Part 2 discusses the typical trading strategies of HFT and the concerns related to those. Part 3 addresses the question why speed is so important in (modern) financial markets and finally part 4 provides empirical evidence on the behavior of HFT. This is based on Eurex data, the largest European derivatives marketplace.

12.1 Definition of HFT and Differentiation to Algorithmic Trading

Despite the attention HFT has received by the public, law makers, and regulators, a common definition of HFT is yet to be found. This irony is best explained by starting with a look in the history.

As long as financial markets exist, the speed of receiving information and executing the resulting actions out of this information has always been a critical success factor for a significant part of the market ecosystem. Early examples include the use of pigeons by Paul Julius Reuters to transmit important stock news from the Paris Stock Exchange to Brussels or Aachen in 1850. Using carrier pigeons to relate messages between the two cities, he bridged the missing telegraph route between the terminal points of the German and the French-Belgian telegraph lines. His idea saved hours. Another example is the Chappe Telegraph. This system of communication relays, a precursor to the modern telegraph, was designed by a Frenchman named Claude Chappe. Each line consisted of signal towers built every 10–20 miles and operators in each tower kept their eye on the adjacent towers through a telescope. Using semaphore signals, they could send messages at what was then considered a staggering speed. Furthermore, the pneumatic tube system of the New York Stock Exchange (NYSE) was launched around 1930. With this, the NYSE went to a great length to ensure that speed location differences within the building did not matter, a principle which still applies in modern financial markets. As a matter of fact, this contends that the speed-sensitive exchange participants try to be in the co-location center of the exchange (comparable to the NSYE building in the 1930s) and the exchanges make sure through the same cable length and identical gear that everybody in the center is treated equally, as the pneumatic tube system did. Obviously participants who do not use the co-location center have a speed disadvantage but this affected also those who were unable to get an office in the NYSE building. The only difference is that today the space in the co-location center is virtually unlimited while in the old days, the space on the floor and in the exchange building was obviously strictly limited.

Against that background, HFT seems to be a natural evolution, resulting from two forces: fierce competition between market participants and technology advances which is used as a competitive element. A prerequisite for HFT and algorithmic trading in general has therefore been the exchanges and marketplaces being transformed into electronic venues. The term "HFT" has been introduced around 2006, but there has been no single event which could be seen as a starting point of HFT. For example, on Eurex and its predecessors DTB and Soffex, in the 1990s, most proprietary futures trading was done manually while in options the market makers were forced from the beginning to be competitive on speed through electronic means. Today these strategies are seen as HFT, while in those days the term did not exist yet.

12.1.1 General HFT Definition

In general HFT is a technology used to implement a wide variety of trading strategies; most of these existed for many years. There are two principal ways to define HFT further; both are discussed here: a qualitative-descriptive definition as well as a mathematical-technical definition. However, as HFT is a technology, it is not possible to have a 100 % clear definition of what activities or trading desks should be considered HFT and which not.

To start with the qualitative/descriptive approach, HFT is obviously a type of algorithmic trading but it needs to be differentiated from the algorithmic trading executed by institutional investors and brokers/banks acting on behalf of these participants.

Common factor of all algorithmic trading is that in general a computer generates orders without human interaction as it implements predefined and pre-parameterized trading strategies. Algorithms employed by institutional investors typically have the intention to minimize market impact of large orders by working those over time and various venues. The resulting positions are held for a relatively long period, i.e., weeks, months, or even years.

In contrast, HFT is typically characterized by trading for their own account; the ability to add, modify, and delete orders within very short time periods (Milliseconds); and the holding of positions for short (intraday) time periods. The HFT activity will be based on a latency-minimizing trading infrastructure.

The problem with the qualitative/descriptive definitions is that it is not possible to define it in a way, whereby HFT firms are characterized by criteria which all of them fulfill and at the same time these criteria do not apply to others. To make a simple example: A hedge-fund dealing on behalf of their funds trading global macro will be a very good fit to the institutional investor algorithm criteria, but from the moment that it starts to also do short-term arbitrage, it will fulfill most of the HFT criteria.

Looking from an exchange point of view at this definition problem, the exchange has the "know-your-customer" advantage. Therefore exchanges are able to do a judgment call based on customer relationships but also the analysis of the activity of the members. This in itself is obviously not a sound basis for any academic or objective research, but this "know-your-customer" information can be used to validate objective HFT definitions.

As a consequence Eurex tried to develop an objective mathematical-technical approach on how to measure HFT behavior and validated the results against the available “know-your-customer” information.

12.1.2 A Mathematical-Technical HFT Definition

The cornerstone of this approach is the assumption that all HFT firms have a high dependency between their profitability and their latency pattern when executing orders at exchanges.

In contrast to following the beaten path in trying to define HFT behavior using criteria like number of trades, overnight positions as close to flat as possible, mean reversion of positions, numerous short-lived orders with follow-up cancellations, etc. this approach tries to use the unbiased measurable latency sensitivity of exchange participants using their transactions arriving at the exchange trading system level in comparison to competitors.

In theory, transaction arrival at the exchange level can be predicted as long as it is uncorrelated. The probability of such transactions arriving with a time difference of t is given by the following formula, where μ is defined as the mean arrival rate of incoming transactions with t being a time interval:

$$f(t) = 1/\mu e^{[(-t)/\mu]}$$

Using this formula we can project the expected theoretical inter-arrival distribution and compare the result with the observation from reality for each and every trading participant.

Pretending that transaction arrival at the exchange is uncorrelated would result in the same distribution of intervals for all members (Fig. 12.1).

To simulate this, we generate a large amount of random numbers between zero and a billion. Assuming those to be time stamps of transactions arriving in our trading system we sort them in ascending order and calculate the difference between two consecutive transactions. Next, we count the occurrences of inter-arrival time interval and plot the respective chart (Fig. 12.2).

The chart depicts the number of observation for any given time interval between two consecutive messages. Taking the log of this function will put the context into a linear relation (Fig. 12.3).

Looking at the transaction arrival data from our trading system, we understand that trading is indeed correlated since we notice a massive burst of transactions around specific points of time (Fig. 12.4).

This effect can easily be explained, as all participants with a latency-sensitive trading pattern will react on the same given signal and thereby increase the number of observations of short time intervals. It is important to emphasize that we do not focus on the event itself but only on the reaction of the participants in relation to each other (Figs. 12.5 and 12.6).

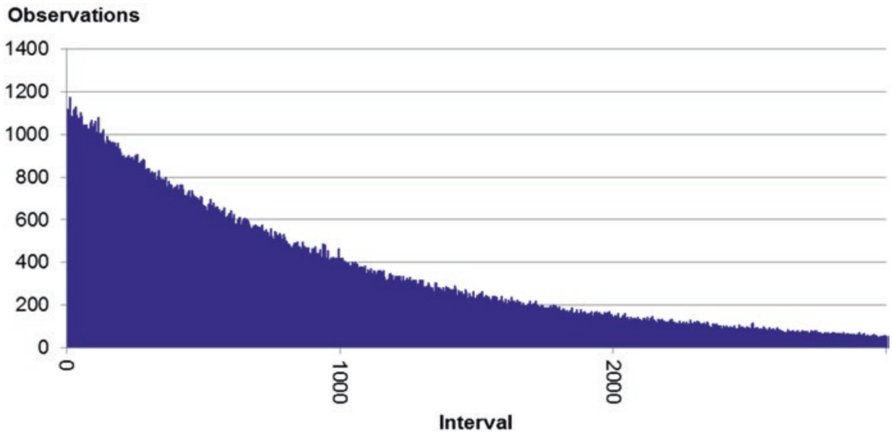


Fig. 12.1 Stylized uncorrelated transactions

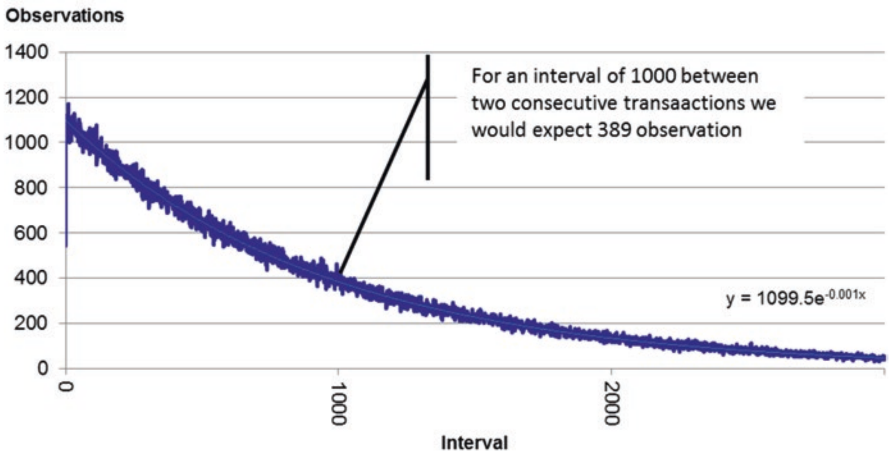


Fig. 12.2 Frequency distribution of intervals

To gain insight, we identify and analyze correlated transaction arrival in our trading system in order to define the latency sensitivity of our members. This can even be enhanced to a more granular view based on the technical connection used to send transactions or on the person responsible for the transaction, the trader.

For consistency in this approach we omit multiple consecutive transactions from the same trading participant and take those ones into consideration which come from different trading participants. As a result of this theoretical approach combined with the actual observations from our production data, we expected that the frequency of lower intervals is higher than the theoretical values for a random distribution just because of the correlated transactions.

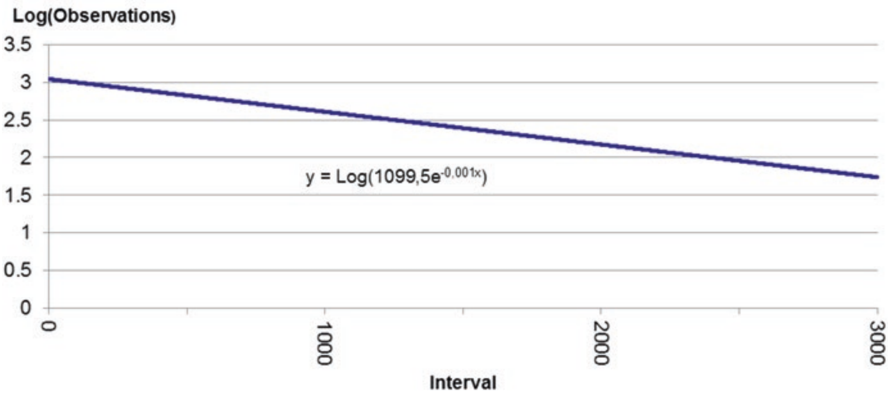


Fig. 12.3 Example of linear relation

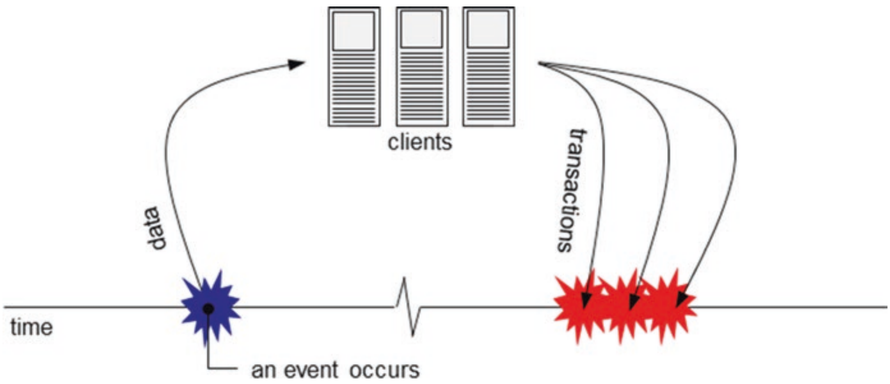


Fig. 12.4 Transaction burst

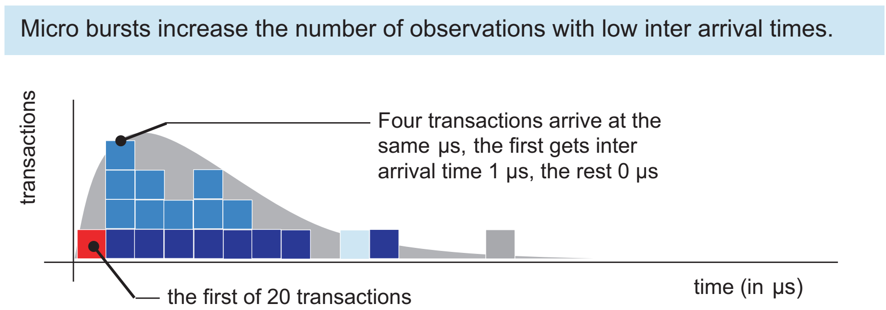


Fig. 12.5 Micro burst

Fig. 12.6 Micro burst frequency distribution

t	E(X)	X
0	3	9
1	2	8
2	1	1
3	0	0
4	1	1

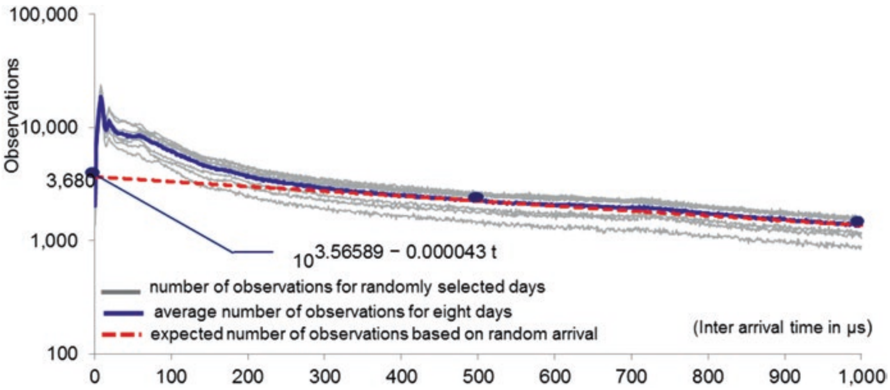


Fig. 12.7 Inter-arrival time frequency distributions

The graph below depicts the actual inter-arrival time frequency distributions for eight randomly selected days in November and December 2012. Due to the log scale of the observation axis, the relationship seems linear (Fig. 12.7).

Focusing on the interval above 500 µs, we observed the expected exponential relationship. Based on our described random process which provides a very good fit for the higher intervals, we notice up to four times (at 8 µs) more observations than expected for random arrival for lower inter-arrival times (<500 µs). Of all observations within the first millisecond above the predicted level from our random process, approximately 86 % occur in the <200 µs area and approximately 67 % occur in the <100 µs area. We therefore presume that most of the nonrandom arrivals are near-simultaneous reactions from strategies using HFT techniques to market data (Fig. 12.8).

Using this methodology to determine the nonrandom part of participants' transactions gives us an indication about their latency sensitivity, and hence their HFT-ness. As the measured dimension is not unswayable, this methodology only contains a decision that is subjective: whether the borderline between excess and conformity is expected with respect to the values from random distribution.

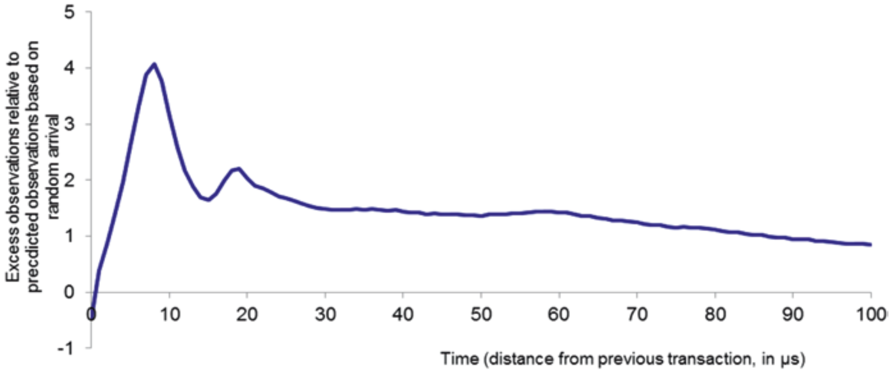


Fig. 12.8 Excess oberservations relative to ramdom arrivial

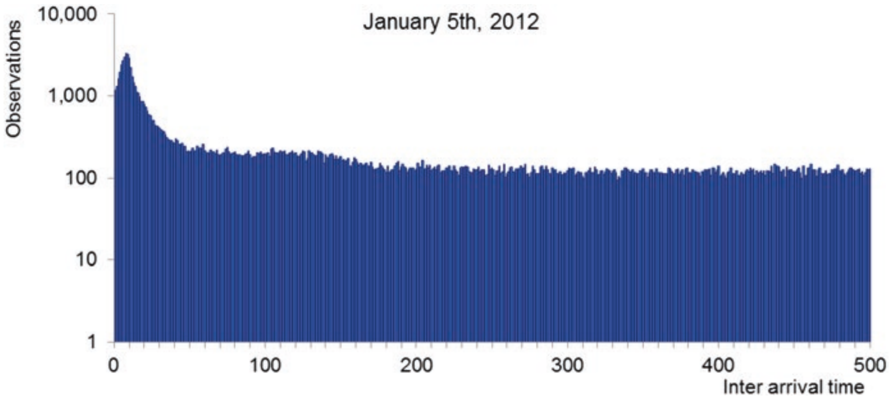


Fig. 12.9 Example accounting for a latency-sensitive participant

Here is an example accounting for a latency-sensitive participant (Fig. 12.9):

Whereas this data accounts for a clearly latency-insensitive participant (Fig. 12.10):

Using this method on actual trading data, we deliver a list of members with excess in the short intervals (see Fig. 12.11), the constituents of which are our well-known HFT participants that could also be on our list by just using the know-your-customer principle. The methodology proved to be an easy flash test to find new latency-sensitive participants.

The table only shows the relation between the intervals of 0–10 and 0–1000 μs. Typically we would expect approx. 1 % to be normal; everything above is "excess" and is deemed to be a hint on HFT activity.

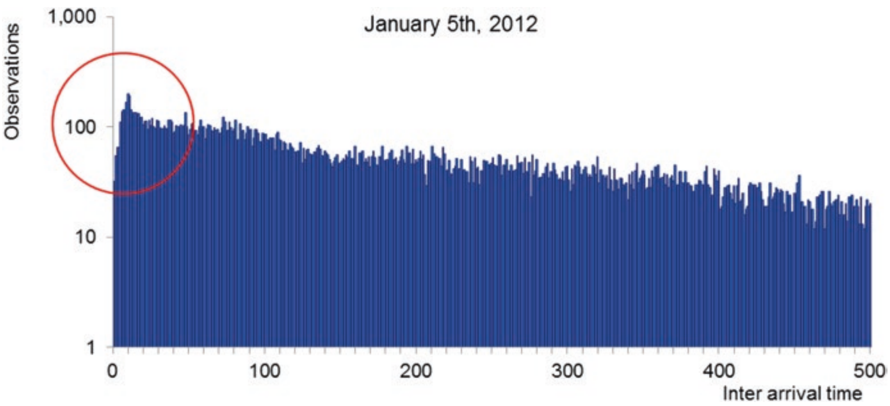


Fig. 12.10 Example accounting for a latency-insensitive participant

Inter arrival time pattern comparison - 5 January 2012				
#	Member	0 µs - 10 µs	0 µs - 1000 µs	Ratio
1	AAA	13.374	79.517	17%
2	BBB	20.491	125.464	16%
3	CCC	24.111	166.402	14%
4	DDD	3.325	25.863	13%
5	EEE	1.230	11.307	11%
6	FFF	8.081	78.635	10%
7	GGG	1.371	19.704	7%
8	HHH	1.842	29.208	6%
9	III	1.660	26.607	6%
10	JJJ	1.496	24.383	6%
⋮	⋮	⋮	⋮	⋮
32	MMM	1.143	43.665	3%
33	NNN	35	1.416	2%
34	OOO	163	6.597	2%
⋮	⋮	⋮	⋮	⋮

Fig. 12.11 Inter-arrival time pattern comparison (example)

12.2 HFT Trading Activities

Trading strategies using HFT techniques can generally be grouped into four categories.

12.2.1 Market Making/Liquidity Provision

A liquidity provider typically contributes two-sided orders, i.e., quotes to markets in order to earn money from the implied bid-offer spread. Typically the provider does not have a preference for one side of the order book/market.

No efficient continuously trading financial market can exist without some participants acting as liquidity provider/market makers. That means even highly liquid futures, where exchanges have usually no dedicated market making schemes, can only be robust to shocks if there are participants who act as liquidity providers.

In modern electronic markets, it is effectively not possible, to provide liquidity, without utilizing HFT technology. The reason is that liquidity providers generate quotes on a certain time-sensitive information basis. These quotes are passive, i.e., can be traded against by everybody. If the underlying information for the active quotes changes, the liquidity provider quotes are still in the market even though they are outdated. Accordingly, the liquidity provider needs to update its quote as swiftly as possible in order to avoid to be taken advantage of at its outdated prices (adverse selection).

In summary, one can implement HFT technology on non-liquidity-providing strategies (see below) but it is highly unlikely to be a liquidity provider in modern electronic markets without being seen as a low-latency trader conducting HFT. As a result, a significant portion of HFT activity is related to liquidity provision. Based on the criteria stipulated by BaFin for qualifying as HFT, we can see approximately 95 % of all transactions coming from HFT participants in the liquidity provision area. The attached chart shows to development over time (Fig. 12.12).

12.2.2 Arbitrage

Arbitrage strategies seek to monetize price differences between identical or related instruments. Those price differences are usually short lived as they are removed by such arbitrage strategies. A classical arbitrage example would be that a stock trades at different venues at different prices. Those price differences are swiftly removed by the arbitrageurs selling the expensive stock and buying the cheap stock. An example of arbitrage between nonidentical instruments is statistical arbitrage whereby there is a statistical mispricing of one or more assets based on the expected value of these assets; that is, assets should stay, based on statistical analysis, in a certain price relationship.

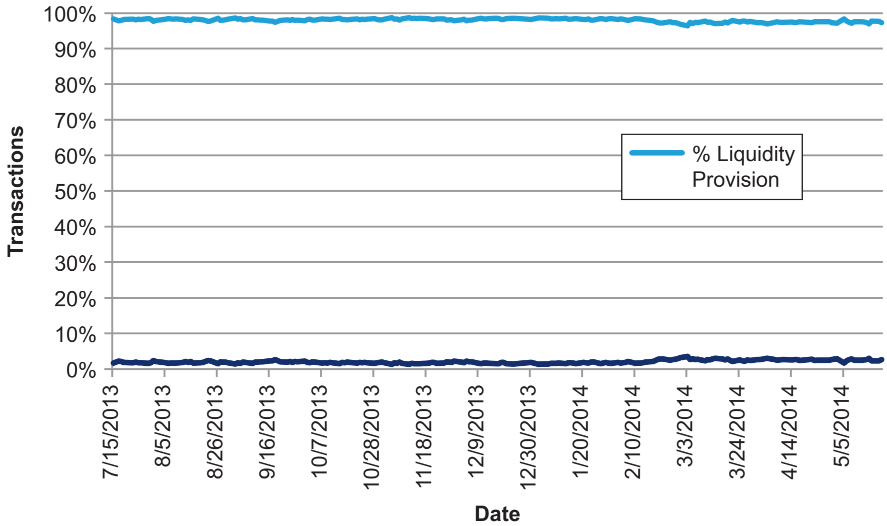


Fig. 12.12 Distribution of HFT members (BaFin criteria)

As the arbitrage opportunities are only short lived in modern financial markets, taking advantage of this requires HFT technology. Arbitrage is beneficial to the market as the prices across different venues and/or related products are on a highly competitive arbitrage-free level reducing the need to compare prices at different venues for the investing public.

12.2.3 News Trading

Unexpected news typically cause prices to move. The relative small and specialized HFT news trading community seeks to benefit from this by being the first market participant to digest and respond to usually prescheduled news. Traders who want to benefit from being able to react to news first not only need to be able to act very fast but also need to be able to understand market sediment. For example, if an unemployment-related figure is newly reported to be worse the market reaction will depend on the ex ante expectation of the market. So even in case of a worse unemployment figure, the market might go up due to the fact that the market expected an even worse number.

12.2.4 Liquidity Detection Strategies

Liquidity detection strategies are thought to be controversial by the public and also by some institutional investors. To fully comprehend the issue, one needs to dive a little bit into the market structure.

Besides news, the driver for price shifts is large institutional orders; for example, a mutual fund decides to build a position in a certain stock whose additional demand will increase the stock price. The mutual funds' decision will have a market impact leading to a conflict between the liquidity provider (now HFT) and the institutional investor who will carry the market impact cost. This conflict is as old as there is a market price-building mechanism. To make a concrete example, we assume that the fair value of a stock is ten monetary units. The liquidity providers' offers account for 10.05. The new fair value after the institutional investor placed its large demand is at 10.20. Obviously, the institutional investor would like to get all his or her stock at 10.05. He tries to do that by taking advantage of the liquidity provided across various market venues. If the institutional investor succeeds, it still will push the fair value to 10.20, as the liquidity providers need to buy the stock back after acquiring a large short position from the institutional investor. In this case the market impact of 0.15 per stock would be completely carried by the liquidity providers, i.e., HFT, thus creating a loss for them.

Therefore, liquidity providers use statistical models to detect patterns in order to make likelihood-based calls, on where the institutional flow is going. In our example that means when the institutional investor starts buying at 10.05, it is likely that at other markets the offers will change towards 10.10 when the buying pressure continues towards 10.20 or even 10.25. The result is that the institutional investor needs to carry a large portion of its market impact on its own. This is normal but it generates frustration when the institutional investor sees at the beginning a much bigger quantity displayed at the 10.05 offers across markets than he or she is able to get at that price.

Consequently, most liquidity providers use liquidity detection strategies as a defensive measure. However, in competitive electronic markets, there are also HFT firms which used the liquidity detection strategies to parallelly run with the institutional flow, i.e., buying when their statistical models indicate that there is institutional buying pressure. As a result the fair price moves short term not only to 10.20 but also to 10.25 or even 10.30 due to the additional buying power.

12.3 The Importance of Speed in Modern Financial Markets

As outlined above, electronic financial markets cannot operate without liquidity providers and those firms need to employ HFT technology. The biggest risk of liquidity providers is that they are not able to update their quotes or orders swiftly enough when new information arrives. Updating the quote/order has three components:

- Receive the new information
- Calculate the new prices/order parameters
- Replace the outdated quotes at the market venue

Accordingly, liquidity providers need to invest in technology/speed in all of those three dimensions: The faster the liquidity provider, the smaller the risk, and the higher its liquidity contribution.

A minimum order resting time will inhibit the liquidity provider in the third dimension. It will create a situation in which the liquidity provider is unable to remove its outdated prices while aggressing strategies can take advantage of the outdated prices. As an immediate consequence, the liquidity provider will either massively increase its spread or leave the market completely.

The long-term consequence is always the same: The market gets uncompetitive and will move in a different jurisdiction.

An internal Eurex study completed in 2012, when the discussions around order resting times (ORT) started, can give insight into this effect and predict the increase of spreads caused by a potential order resting time based on empirical data.

For analyzing the reluctance of liquidity providers to take volatility-based risk we took several days from August 2012 and calculated the average spread quoted for the front month of the DAX future (FDAX).

In order to predict the impact on spreads caused by a potential ORT we needed to analyze the relation between the probability that prices move before participants can update their orders (volatility) and the inherent compensation for risks taken when providing the visible liquidity (spread). To get the most granular insight into the risk aversion we calculated the standard deviation of the mid-price between the best available bid and ask prices from 1 ms to the next (Fig. 12.13).

The graph shows ten daily observations of the spread and volatility on a 1 ms basis. The green dot represents a high-volatility day, August 12th, 2012. We can derive from the chart that the spread increases with the standard deviation by a factor of ten and changes in volatility explain 60 % of changes in spreads. To gather information about the volatility component we measure the standard deviation for several fixed time frames; 50, 100, 250, 500, and 1000 ms.

Fig. 12.14 depicts how volatility increases on a micro level dependent on the measured time frame.

Each line in the chart represents a day, whereas the dots represent the time frames where volatility was measured.

Pretending that the actual time frame it takes to add or delete an order in our system (approx. 1 ms) is an ORT itself and is a crucial part of the above-shown figures in combination with the assumption that the risk aversion of liquidity providers stays unchanged, we are able to predict spreads for different order resting times (Fig. 12.15).

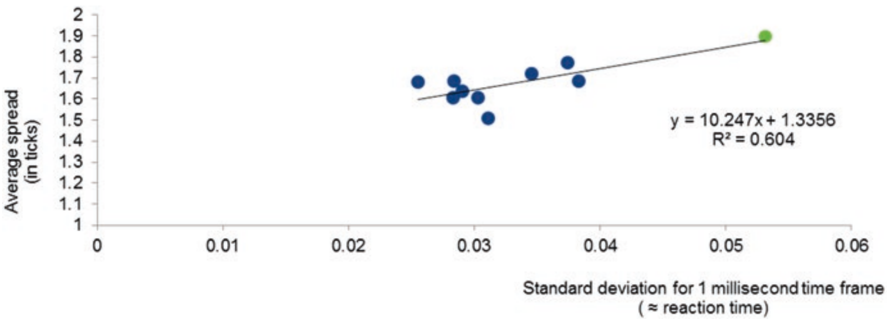


Fig. 12.13 Overview of daily observations of spread and volatility on millisecond basis

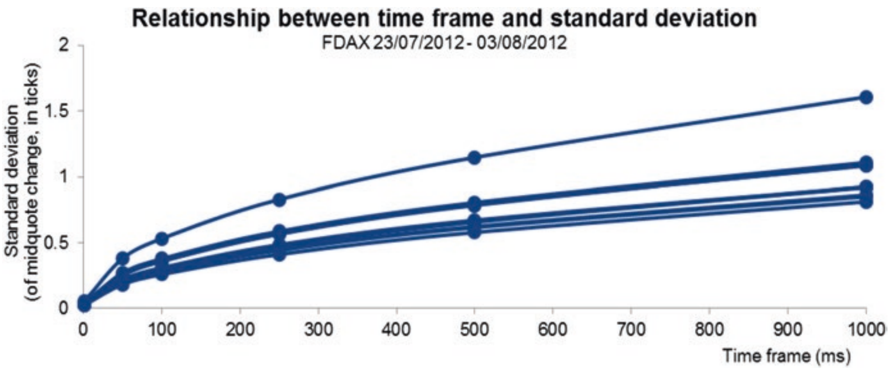


Fig. 12.14 Relationship between time frame and standard deviation

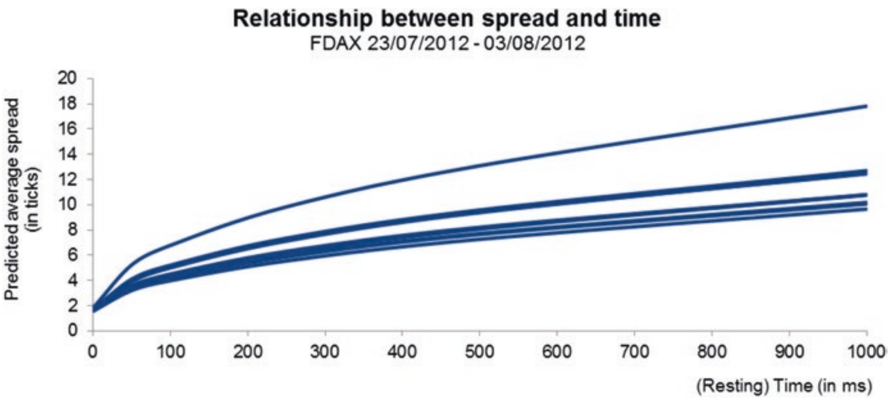


Fig. 12.15 Relationship between spread and time

The graph shows the estimated spread for 10 different days (lines) and several resting times. Eye-catching is the concave positive relationship between minimum order resting time and spread.

Reading the chart, a resting time of, e.g., 500 ms on a normal trading day would lead to a spread widening of 400 %, from 1.5 ticks to 9 ticks. On a very volatile day, a spread widening of 600 %, from 2 ticks to 12 ticks, can be expected.

So in summary, modern electronic markets require participants to provide liquidity. A prerequisite of liquidity provision is to be competitive on speed, as speed is and always has been a predominant competitive element in trading. In other words, no matter how often or seldom the information related to a specific instrument change, if it changes, it takes the fastest and best technology available to be a competitive liquidity provider; these days, using this technology is called HFT.

12.4 Empirical Evidence on the Behavior of HFT

Market quality as such can be defined as a function of spread width and book depth. The Eurex Liquidity Measure (ELM) is ideal to get a first indication about those components. The ELM measures the round-trip market impact cost of somebody executing a €10 million market order against the public order book. It consists of two components: The liquidity premium (LP) measures the spread cost of a simple 1-lot round-trip market order, and the advanced price movement (APM), which measures the additional market impact cost when a €10 million market order is executed in a round trip via market orders, in the DAX future (FDAX).

It is therefore mirroring the displayed size in the order book: The larger the price impact, as measured by the ELM, the smaller the available size in the order book. At times of crisis, the market impact cost increases, as participants scale down their risk profile, implying somewhat wider spreads and significantly reduced sizes. In general the liquidity readily available in the order book is slightly worse in Q3 2012 than in 2005 (Fig. 12.16).

This is an effect of a change in market behavior. A major driver for this is the use of execution algorithms by the buy side, which has vastly reduced the placement of large resting orders by the buy side in the transparent order book.

HFT adds significant liquidity, but their order sizes are typically smaller even though the orders are faster compared to other market participants. This ensures that participants get optimal execution even on a microsecond scale.

The ELM is because of the change in market behavior, a suboptimal indicator, and we will focus on spread resilience instead, i.e., how fast the spread between bid and ask recovers after a large trade hit the order book. For a more precise view on the topic we separated the analysis into two parts:

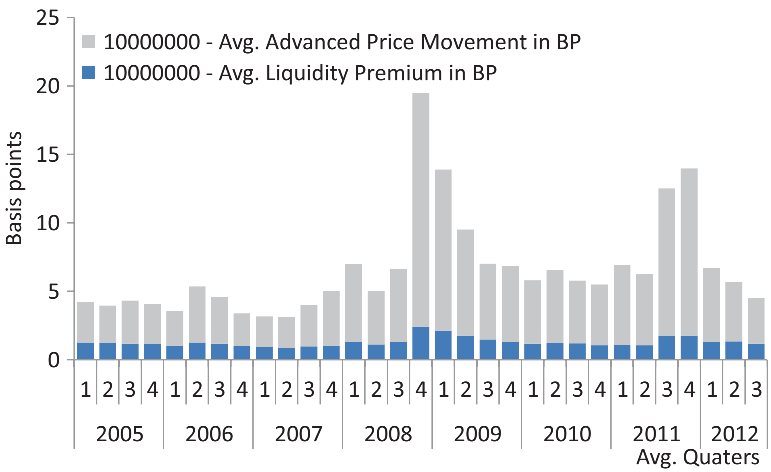


Fig. 12.16 Example of the Eurex liquidity measure

Order book liquidity share of low-latency participants at the "best bid offer" (BBO) during average-day trading

At first, we have to define how we quantify the resilience. Therefore we calculate the average traded size of a product (here: FDAX front month). We assume that trade sizes of at least ten times the median will have enough impact to move bid or ask; on the other hand we ensure using that number that we have enough samples to check the quality of our results. Fig. 12.17 shows a stylized picture of the expected market behavior before and after a large buy order hit the order book.

Focusing on timescales, where a human interaction is nearly impossible (<200 ms), and using data from 2010 to 2012, the result shows a much faster return to former levels in the spread in 2012 compared to 2010 after a hit (Fig. 12.18).

Compared to 2010, the liquidity in the DAX futures (FDAX) became much more resilient. The averages of 2010 and 2012 converge around 500 ms after a big trade.

As we see a lot of trading activity on the back of large trades, the faster return to smaller spreads increases the quality of the executions of those related trades. Taking into consideration that nearly no human interaction can take place in such short timescales, it seems justifiable to attribute that positive aspect to low-latency participants.

Order book liquidity share of low-latency participants at BBO in crisis times

Even if the public perception is that low-latency participants might provide liquidity during normal market circumstances, the majority highly doubt that those participants are also providing liquidity during crisis times. Market turbulence in combination with exceptional high volatility levels often puts low-latency trading participants in the spotlight. It is publicly assumed that price volatility would be significantly reduced if high-speed trading did not exist.

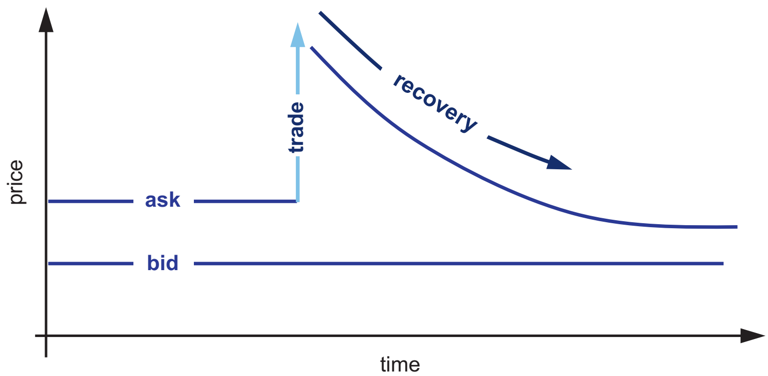


Fig. 12.17 Example for expected market behavior before and after a large buy order

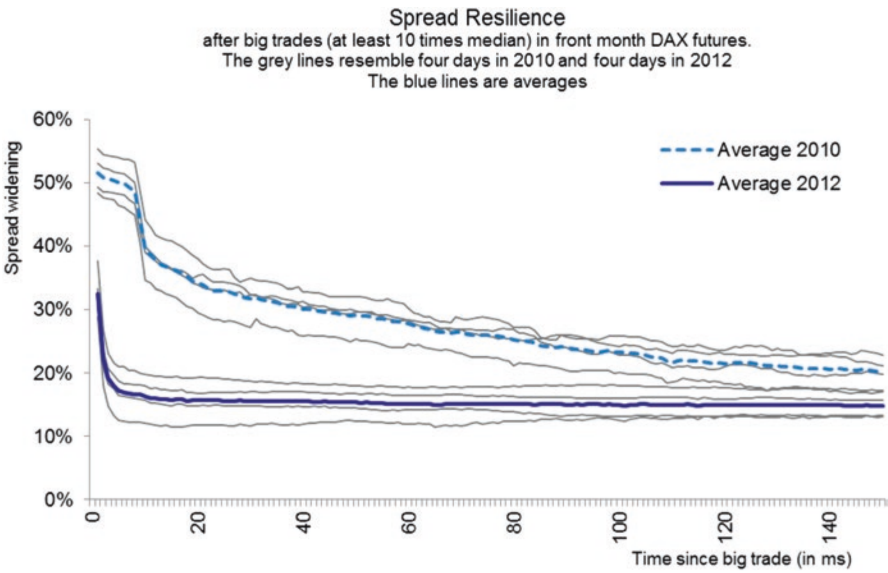


Fig. 12.18 Example for spread resilience after big trades

To verify or falsify this allegation, Eurex investigated some significant market situations where the products in scope moved a huge percentage and especially examined the participation of low-latency participants during the event and in the immediate aftermath of such an event.

12.4.1 August 25th, 2011 Futures on DAX (FDAX)

In the afternoon of August 25th, 2011, the FDAX lost more than 4 % of its value within 17 min, only to reverse this move by 2 % within minutes. The decline was caused by a big institutional order, which was sliced and diced by algorithms into a large number of smaller sell orders flooding the market during that relevant period of time. The total amount of the sliced and diced orders sent was 6000 contracts.

At the starting point, the order book was highly liquid with an average volume per minute of slightly above the monthly traded minute average of 300 contracts and around 60 members active on the bid and ask side of the order book.

During the peak minute at 16:02 a high number of small orders were processed with only small price increments, causing a peak turnover of 4700 contracts during that particular minute. The number of participating members during that minute doubled (Fig. 12.19).

The high number of members involved on both sides of the market during this event shows the high variety of trading interests in our markets, a key driver for liquidity and quality.

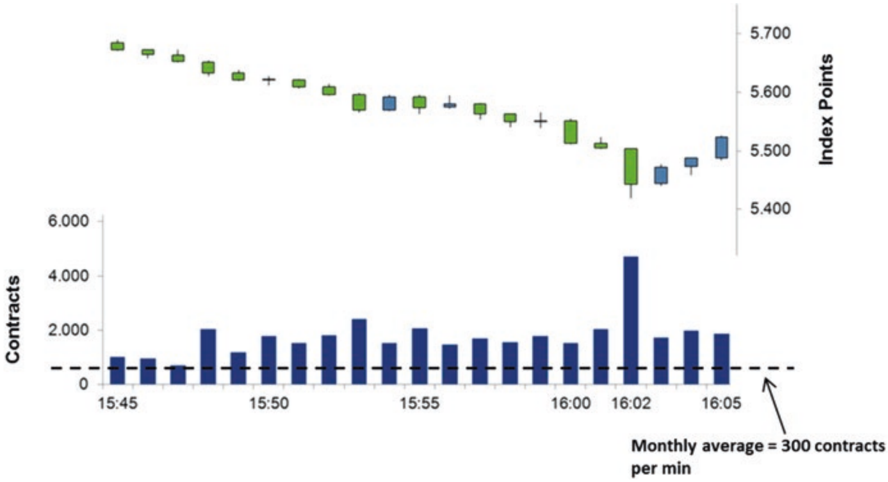


Fig. 12.19 Development of Futures on DAX (FDAX) on August 25th, 2011

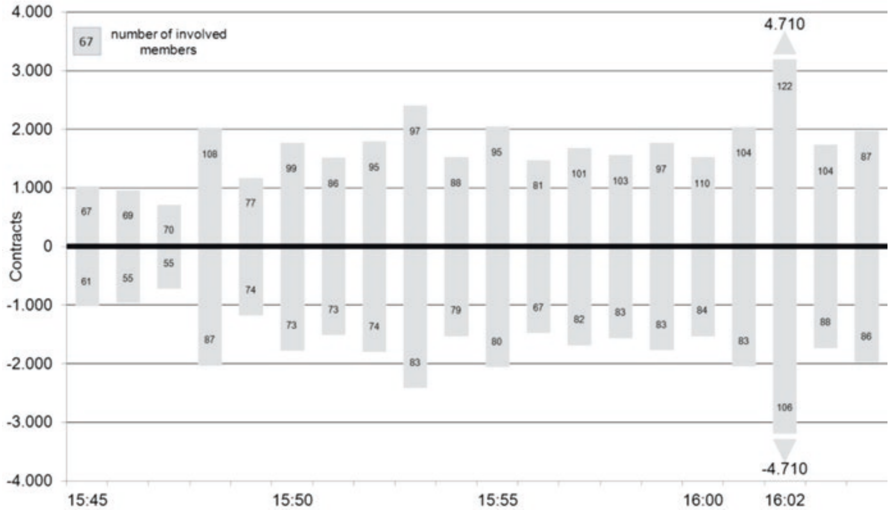


Fig. 12.20 Involved members in Futures on DAX (FDAX) on August 25th, 2011

The total of around 200 different trading members acted during the time slice in scope as buyers in a falling market, including but not limited to low-latency participants (Fig. 12.20).

A large junk of the enormous liquidity was provided by low-latency participants applying liquidity provision and arbitrage strategies. The often-heard allegation that the strong movements are accelerated by computer-based trading strategies of low-latency participants cashing in by simply using their speed could not be observed. For more details and an insight on market activity during volatile periods of trading see our homepage at http://www.eurexchange.com/exchange-en/technology/high-frequency_trading/ with videos on low-latency trading activity.

12.4.2 April 6th, 2014, Futures on DAX (FDAX)

On 6 February 2014 at 13:45 CET the ECB was scheduled to publish its announcement on interest rates. The publication of the rate decision followed the standard ECB protocol and was in line with market expectations. As scheduled ECB decisions always have a potential to move prices/markets; typically the order books and trading as such tend to get thinner the closer the deadline for the announcements gets. This is in anticipation of the potentially market-moving information and is typically adjusted back to normal as soon as the information is released. On the particular day in scope, just 4 s and 403 ms after the ECB's announcement on interest rates was made public, a strong selling pressure emerged in the FDAX in form of sell orders. In the following 414 ms those orders started to push prices sharply lower while a total of 49 sellers and 82 buyers traded 1488 contracts (Fig. 12.21).

Such a situation can arise from the fact that one or more participants are placing one or some large orders in the order book to adjust their respective position in relation to the just published information. In cases where the size of the orders surmounts the available, still decreased liquidity in the order book, such an order entry, can cause a move in prices. Where under normal circumstances this happens only very seldom, it still might happen at any time during trading. Accordingly, the Eurex T7 trading system has built-in functionality to safeguard and manage the

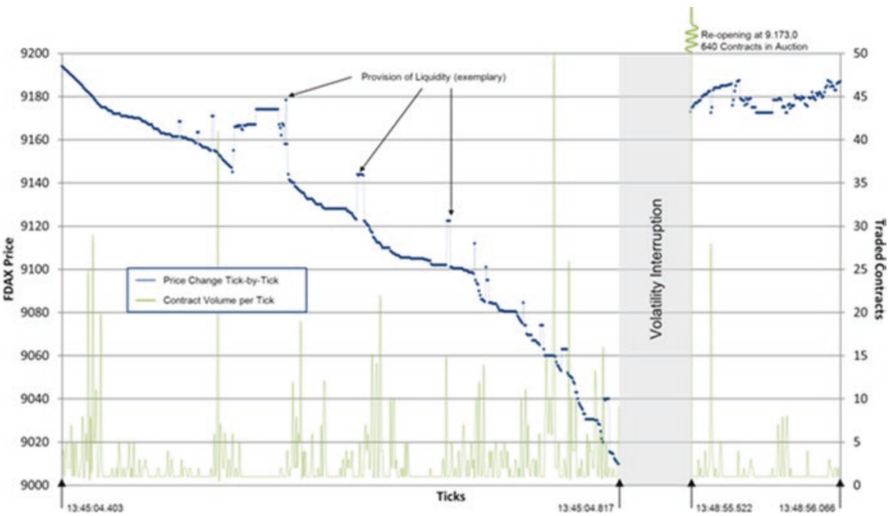


Fig. 12.21 Price development and traded volume in FDAX on February 6th, 2014

impact of such events. Also in this event those safeguards worked out and halted the market to guarantee fair and orderly market conditions and executions.

The up spikes in the chart show clearly that new orders, coming into the order book, provided new liquidity at a better level and market orders resting in the book, because of our market order matching range that halted the execution of those due to the missing price references, were executed at better levels than it would have been possible before.

To stress the fact that we talk about a time slice of 414 ms, we tend to attribute at least the newly provided liquidity again to the low-latency participants, as a human reaction in the time interval is at least very unlikely.

12.4.3 Effects of Low-Latency Participants Engaging in New Products

As already stated, low-latency participants play an important role when it drills down to liquidity provision and market depth (order book elasticity). Their ability to digest news and market information at high speed allows them to be in or back in the book faster than anyone else. This increases the quality of executions after the release of news or big orders hit the book and caused price moves.

Especially in new products, where the liquidity is still growing, low-latency participants can be extremely helpful and important. In order to gain insight into the importance and the effects of low-latency participants engaging in new products we researched the Futures on Italian (FBTP) and French Government bonds (FOAT) traded at Eurex. The products were introduced on Sept 14th, 2009 (FBTP), and April 16th, 2012 (FOAT).

Contrary to general market development, OAT and BTP futures performed pretty well in 2012 and gained even more market acceptance. During the month of August 2012, the spread in both products experienced a stellar improvement.

The development did not seem gradual, and all signs were that a structural break around August 17th occurred. This break is most noticeable in the development of the spread quality depicted as the one tick spread percentage of the day in Figs. 12.22 and 12.23.

As an additional hint for a structural break, the size available at the improved spread can be taken, which also increased starting August 17th, 2012 (Fig. 12.24), for the FBTP. While there is a clear signal in the FBTP data, there is no clear picture in the FOAT (Fig. 12.25).

When looking at the average spread in the two products we again got confirmation of a structural break in the FBTP and no signal for such a break in FOAT. Even though the average spread in FOAT is also decreasing over the period in scope it is not clearly attributable to a particular date (Figs. 12.26 and 12.27).

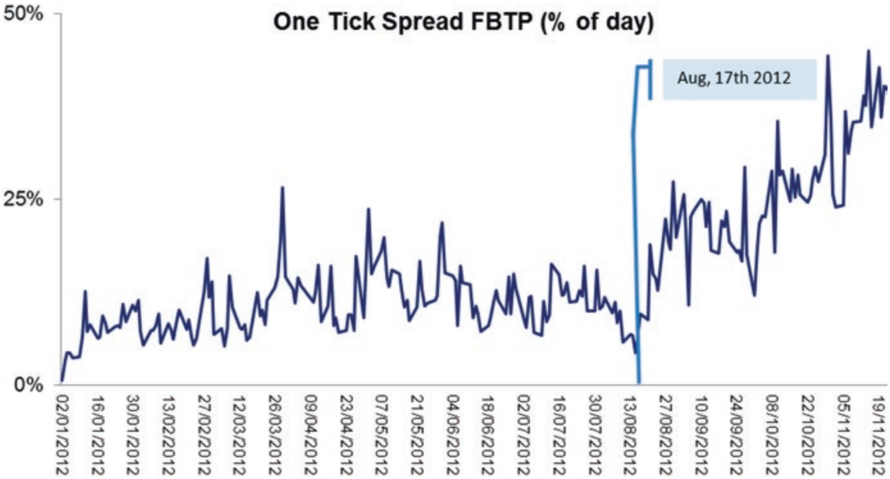


Fig. 12.22 One tick spread in FBTB

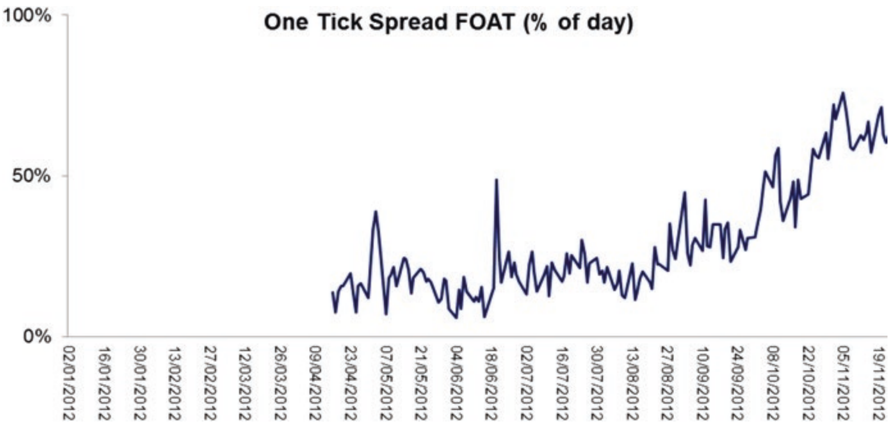


Fig. 12.23 One tick spread in FOAT

A major, identifiable market structural development that took place on July 23rd might at least partially explain the visible change in FBTP markets starting August 17th, 2012.

Here is what happened: up to July 20th, one particular low-latency participant (AAA) provided close to 10 % of the BBO on a daily basis. On the next trading day (July 23rd, 2012) another low-latency participant (BBB) joined in providing the BBO and took over the position of AAA on Aug 17th, 2012, by providing a

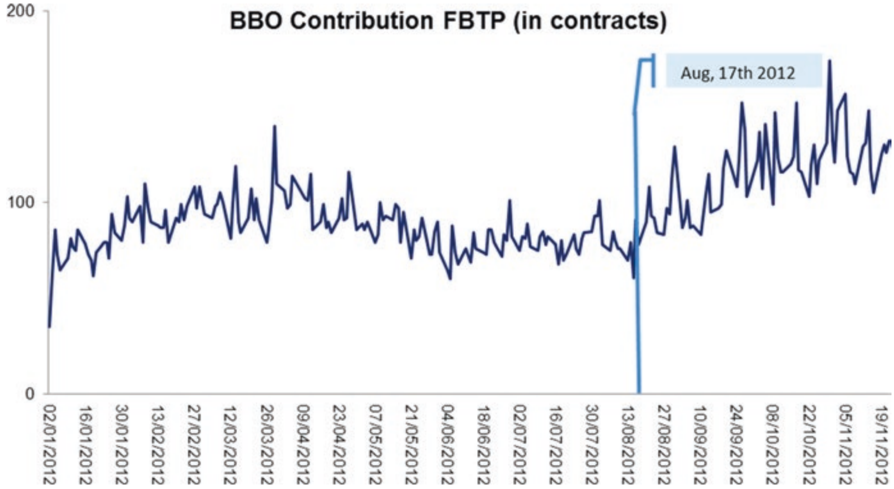


Fig. 12.24 Size available at improved spread for FBTP

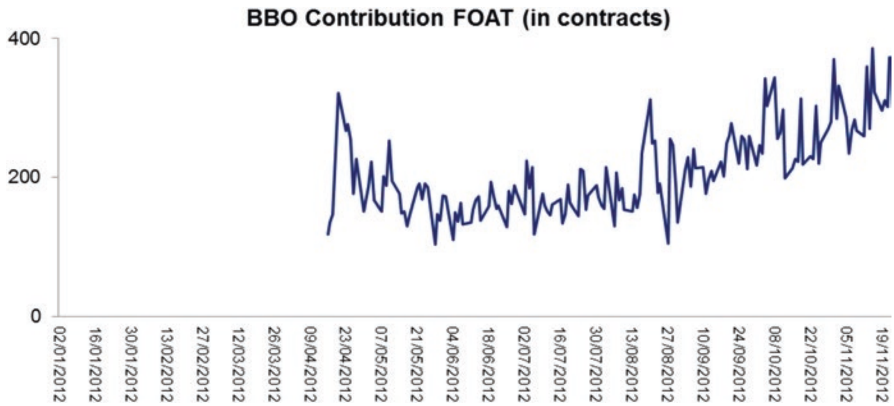


Fig. 12.25 Size available at improved spread for FOAT

larger share of the BBO on a daily basis, while participant AAA left the market over the following 5 days.

Our thesis is that participant BBB's business model is very similar to AAA's, but faster in the execution.

When AAA realized that another participant is running in parallel, but faster in taking decisions and sending orders, AAA specialized in areas of competence with less competition.

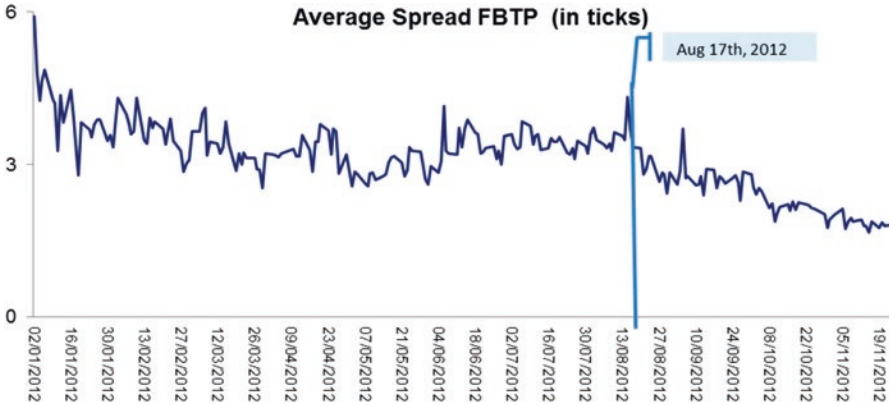


Fig. 12.26 Average spread in FBTP

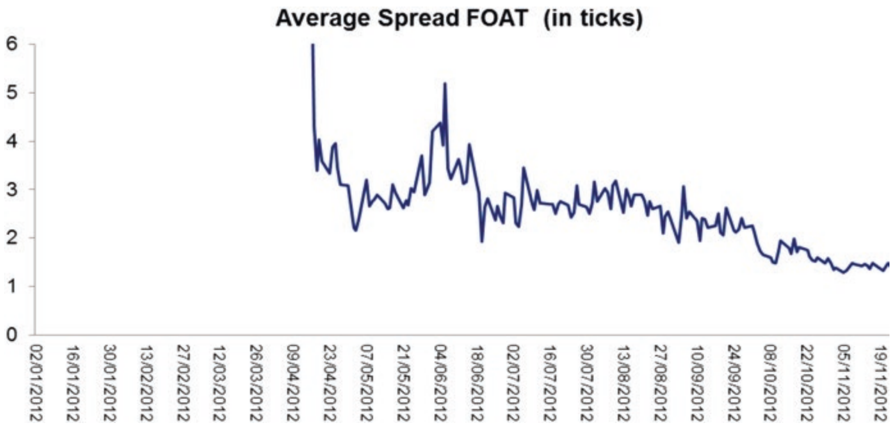


Fig. 12.27 Average spread in FOAT

One piece of evidence in favor of this thesis is the improvement of spread resilience following trades, which is to the benefit of all participants as the executions following relatively large trades would otherwise be suboptimal (at worse prices). A remarkable improvement (red) took place on the date BBB entered the market. This improvement did not take place in OAT futures, where BBB did not enter the market. It shows clearly that the new entry of a low-latency participant into a product can significantly increase the quality of executions (Figs. 12.28 and 12.29).

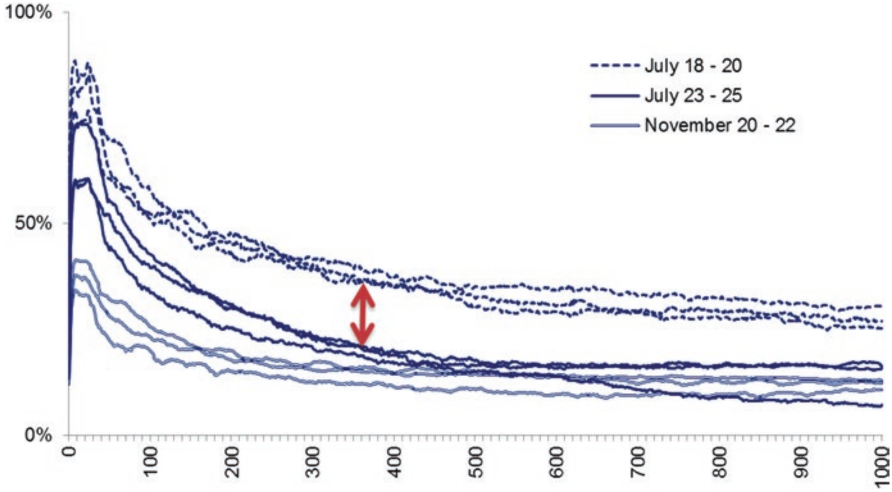


Fig. 12.28 Resilience in FBTP

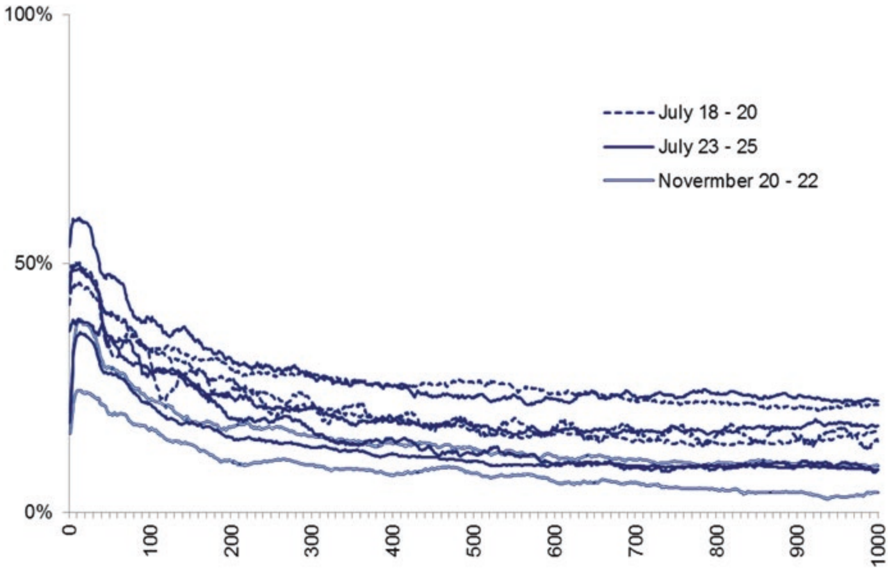


Fig. 12.29 Resilience in FOAT

Bibliography

Bloomberg, Published Article, Apr 2nd, 2014. 'High-Speed Trading Used to Mean Carrier Pigeons. http://www.bloombergview.com/articles/2014-04-02/high-speed-trading-used-to-be-for-the-birds

Hagströmer B. & Norden L. 'The Diversity of High Frequency'. Stockholm University of Business. September 27, 2012.

Lee H. G. & Clark T. H. 'Impacts of the Electronic Marketplace on Transaction Cost and Market Structure' International Journal of Electronic Commerce volume 1, Number 1, fall, 1996, pp. 127–149.

Markham J. W. 'A Financial History of the United States. From JP Morgan to the Institutional Investor (1900–1970)'. Volume II, 2002 M.E. Sharp Inc.

Oxera, May 2011 'Monitoring prices, costs and volumes of trading & post trading services'. Report prepared for the European Commission, DG Internal Market and Services

Reuter, Paul Julius, Baron de. Encyclopedia Britannica (11th ed.). 1911

U.S. Securities and Exchange Commission. March 18, 2014. 'Equity Market Structure Literature Review'. Part II : High Frequency Trading. Staff on the Division of Trading and Markets. HFT Overview, Metrics for Defining HFT, Page 4. Statistics Web F/E

Chapter 13
High Frequency Trading: Market Structure Matters

Reto Francioni and Peter Gomber

13.1 The Evolution of Electronic Equity Trading Along the Value Chain

Although both media and the public seem to discuss the perceived dangers and threats of electronic trading only since the US flash crash in 2010, in reality, the shift towards electronic trading has been a long-lasting evolution. Often, the starting point of electronic trading is said to be the year 1971, when the National Association of Securities Dealers Automated Quotation (Nasdaq) became the first electronic stock market displaying quotes for 2500 over-the-counter securities. A significant migration process from over-the-counter and traditional floor trading to fully electronic markets took place on both sides of the Atlantic between the late 1970s and the mid-1990s. Starting from the electronification of major international exchanges, significant technological innovations emerged that successively walked up the value chain and led to a far-reaching automation of trading processes; first at Sell Side institutions and in a next step by their customers, i.e., Buy Side firms.

In order to increase market transparency, *exchanges* established public electronic limit order books that aggregate and store open limit orders and match executable orders in real time. In contrary to traditional market maker systems, these central limit order books (clobs) enable all investors to enter limit orders and thereby to reduce spreads and improve market depth. Instead of relying on price determination services provided by expensive market intermediaries like specialists or market makers, the matching of

R. Francioni (✉)
Wirtschaftswissenschaftliche Fakultät, University of Basel, Peter-Merian-Weg 6, Basel, 4002 Switzerland
e-mail: Reto.francioni@unibas.ch

P. Gomber
Faculty of Economics and Business Administration, Department of Information Systems, University of Frankfurt, Grüneburgplatz 1, Frankfurt am Main 60323, Germany
e-mail: gomber@wiwi.uni-frankfurt.de

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_13

orders and price determination was performed by central matching algorithms that serve as the heart of market structure. These matching procedures—that are transparently communicated to any market participant as well as to any investor—treat all orders equally (operational fairness)—mostly based on price-time priority allocation. The transparency induced by the introduction of clobs reduces information asymmetry, enhances liquidity, and supports efficient price determination.[1] Electronic trading systems enable decentralized market access, i.e., allow investor to place orders from remote locations and to trade at significantly lower costs compared to physical floor trading.

On the *Sell Side*, the introduction of electronic exchanges also trigged a far-reaching electronification of order handling processes. This included the implementation of real-time automated price observation mechanisms, i.e., electronic eyes, and automated quote machines that support the fulfillment of liquidity providers' obligations by generating quotes based on predefined parameters. In order to reduce costs and to release traders from time-consuming and mostly standardized order execution procedures like vwap-executions, the Sell Side implemented the first trading algorithms for their proprietary executions that used predefined parameters like total order volume, limit, execution period, or execution aggressiveness for order splitting and order timing.

Also *Buy Side* institutions established electronic trading desks successively to connect via electronic means to multiple brokers and liquidity sources. Especially the introduction of the Financial Information eXchange (FIX) Protocol—today the de-facto messaging standard for pre-trade and trade communication—accelerated electronic interchange of trade-related messages between the Buy Side and the Sell Side. Against the background of an increasing pressure to reduce trading costs and to increase execution quality for Buy Side customers, brokers began to provide direct market access (DMA) tools. In a DMA setup, a Buy Side order is not worked or splitted over time and markets by the intermediary, but forwarded directly to the execution venues using the broker's market connectivity and trading infrastructure.[2] The Buy Side institution has no need to become a direct member of the respective market(s) but trades directly on the market by renting the exchange membership of their Sell Side broker.[3]

Order Management Systems (OMS) and Execution Management Systems (EMS) further improved efficiency in internal order handling of the Buy Side as well as the purely electronic interaction of the Buy and the Sell Side by enabling for connectivity, automated routing and integration with confirmation, clearing

[1] See, e.g., Pagano and Roell [14] or Jain [15].

[2] See [16].

[3] In the last 5 years, Sponsored Market Access has emerged as an extension to DMA. Sponsored Market Access enables Buy Side clients with latency sensitive strategies to connect to the market via their broker's membership and identification but omitting the broker's technical infrastructure in order to achieve latency reduction. In this concept, brokers' risk management only relies on automated pre-trade risk checks—e.g., by setting the maximum number of orders in a predefined time window or a maximum order value—that are implemented within the exchange software and administered by the respective broker. Naked access describes a setup without these pre-trade risk checks and therefore enables only for post-trade monitoring. Due to the possible devastating impacts by erroneous orders and orders submitted by flawed algorithms, the SEC banned "naked access" in 2010 and requires all brokers to put in place risk controls and supervisory procedures relating to how they and their customers access the market [17]. In Europe, naked access is also prohibited.

and settlement systems. Most of these OMS/EMS include broker algorithms like vwap, twap, or implementation short fall algorithms that are used by the Buy Side traders for their order executions. While in this setup, the algorithm is implemented, hosted, and operated by the broker, the Buy Side trader is informed on the basic strategy of the algorithm and instead of sending individual orders now is enabled to specify strategies and strategy parameters that are executed by the broker algorithm. Thereby, the Buy Side, i.e., the ultimate owners of the orders and order flows, achieved more direct control over the order routing and execution processes and execution responsibility was shifted—upwards in the value chain—to the Buy Side. In this process, the role of the Sell Side changed from an intermediary that was fully responsible for order execution to a provider of market access and trading technology.

The introduction of Electronic Communication Networks (ECNs) in the mid-1990s in the USA and the new concept of Multilateral Trading Facilities (MTFs) in Europe since 2007 triggered significant competition between exchanges and these new market players that strive to further reduce trading costs and to enhance execution quality for investors leading to a highly fragmented, electronic trading landscape. Smart Order Routing (SOR) services were introduced first in the US market and later in Europe to support order routing in a fragmented market structure. SORs continuously gather real-time data from multiple liquidity pools concerning the available order book situations to optimize order execution by identifying the highest liquidity and optimal price.

Increasing automation, electronic connectivity and increasing competition among traders to exploit profitable market situations as fast as possible in proprietary trading and to maximize order execution performance for customer orders (agent trading) triggered a demand for lower transmission latencies in the process of order submission, order arrival at the trading venues, and confirmation back to the trader. For traders, speed is a central tool for risk management as it has to be assured that the time differences between the transmission of the order book situation form the trading venue to the algorithm in (t_0), the arrival of that data at the algorithm at (t_1), the sending of an order after the computation effort by the algorithm in (t_2), and the arrival of that order at the order book of the trading venue in (t_3) has to be minimized. In case of a long time span between (t_0) and (t_3), there is a significant risk that the intended execution is significantly different from the finally achieved execution due to market movements between (t_0) and (t_3). As physical distance to the exchange contributes significantly to the total roundtrip latency in order execution, co-location and proximity services were introduced to reduce this physical distance between the market participants' servers and the market server to a significant extent. Against this background, the phrase "High Frequency Trading" emerged. However, the terms "Algorithmic Trading" and "High Frequency Trading" were only discussed in traders and market expert circles until 2010 when the May 6th US flash crash churned the public discussion and triggered a heated debate on the benefits and threats of this evolution in electronic trading.

Figure 13.1 sums up how the evolution of trading technology influenced the interaction among market participants along the trading value chain and shows the

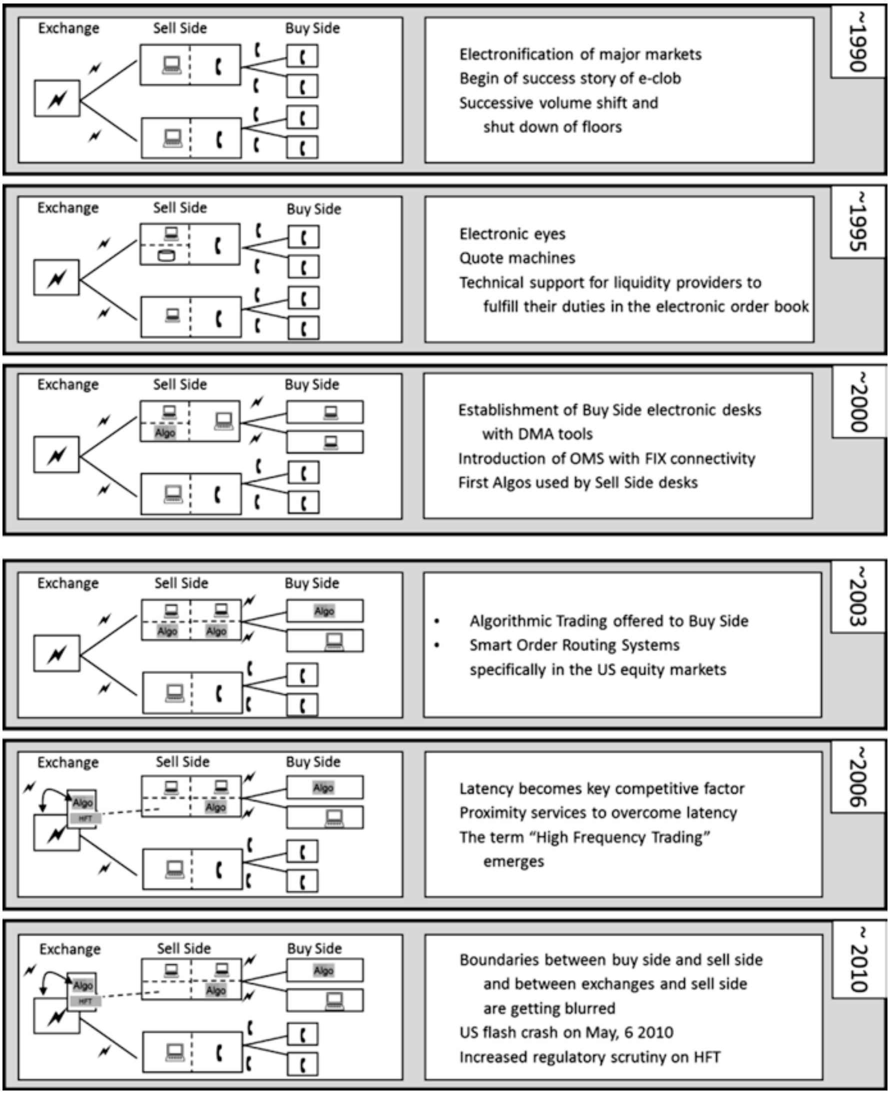


Fig. 13.1 The evolution of electronic trading along the value chain. *Source*: Gomber, Zimmermann [13]

major milestones in the interaction between exchanges, the Sell Side and the Buy Side. The figure also takes into consideration the important fact that while lager Buy Side institutions today are equipped with similar trading tools like their Sell Side counterparts, smaller Buy Side firms still rely on telephone, fax, or e-mail to communicate orders to their brokers.[4]

Based on this review on the evolution of automated trading, the following section will define and delineate the two terms AT and HFT. Based on a comparison of equity market structures in the USA and in Europe, Sect. 13.3 will identify their main commonalities and differences and the role of market structure in electronic trading. Section 13.4 will analyze the major market structure and regulatory changes in response to AT and HFT in the last 5 years on both sides of the Atlantic. Section 13.5 concludes.

13.2 Algorithmic Trading and High Frequency Trading

13.2.1 Definitions and Delineations of Algorithmic Trading and High Frequency Trading

Today, in most industries—among others e-Commerce, automotive or health—the need for a rapid aggregation and examination of huge amounts of data in real time (Big Data) as well as the increasing pressure to reduce cost in the execution of standardized processes has shifted a lot of tasks form humans to software and algorithms. Generally, computer algorithms are defined as an execution of predefined instructions in order to process a given task.[5] In securities trading, the evolution of electronic trading has lead to multiple forms of algorithms that support data aggregation and standard processes in order generation, order routing, and order execution for actors in every stage along the value chain. Figure 13.2 shows the role and positioning of software and algorithms along the value chain.

First of all, algorithms are implemented by trading venues to perform order aggregation, price determination, and order matching at the market back-end. These algorithms, e.g., a matching algorithm of an opening auction, are the technical realizations of an exchange's market model. Secondly, algorithms are used to generate orders by market participants based on predefined parameters either to support efficient execution of customer orders thereby increasing customers' portfolio performance (agent trading) or to implement sophisticated trading strategies that strive to maximize revenues in trading on own account (proprietary trading). Thirdly, algorithms are used in quantitative portfolio management at Buy Side firms to implement

---

[4] In case of retail order flow, the electronification of the Buy Side to Sell Side interaction also refers to retail investors using the World Wide Web to route orders via online-brokerage accounts.

[5] See [16].

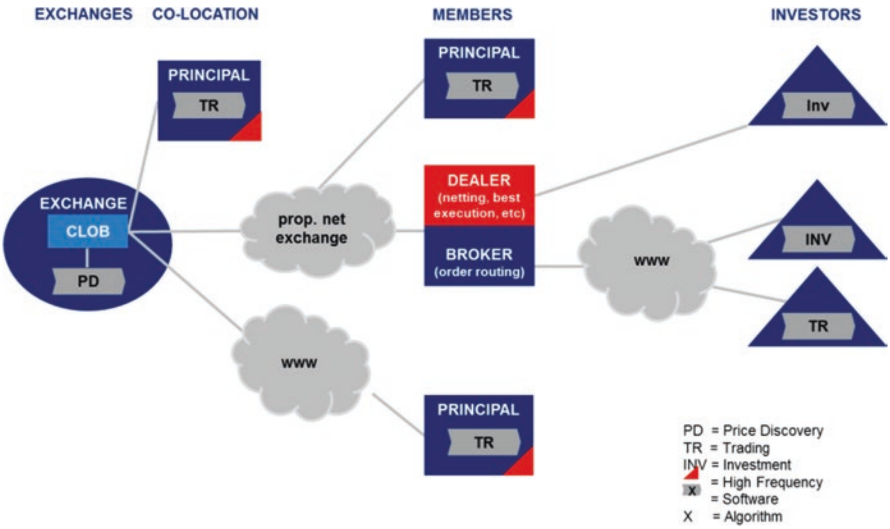


Fig. 13.2 Price discovery and order execution: software to software

mathematical models that, based on statistical calculations and data analysis techniques, predict, e.g., stock returns or stock volatility, to form investment strategies and portfolios.

The terms Algorithmic Trading (AT) and High Frequency Trading (HFT) refer to the second group of algorithms, i.e., these algorithms support market participants by providing a set of instructions on how to generate, process, or modify an order or multiple orders without human intervention.

Despite an intensive discussion on the merits and dangers of trading algorithms, both the academic world and international regulators largely agree on two principles: (1) HFT is seen as subgroup of AT and (2) both HFT and AT are no new trading strategies in itself but technical means to implement established trading strategies by applying sophisticated technology.

In contrary to the wide acceptance of these principles, there is an intensive debate on how to correctly define AT and HFT respectively and how to delineate those definitions exactly.

Academic definitions of AT mainly take a task-based approach and describe the main processes that algorithms perform in order generation and processing, e.g., Hendershott et al. [1]: "*[…] AT, commonly defined as the use of computer algorithms to automatically make certain trading decisions, submit orders, and manage those orders after submission*," Chaboud et al. [2]: "*[…] in algorithmic trading (AT), computers directly interface with trading platforms, placing orders without immediate human intervention. The computers observe market data and possibly other information at very high frequency, and, based on a built-in algorithm, send back trading instructions, often within milliseconds […]*" or Domowitz and Yegerman [3]: "*[…] we generally define algorithmic* *trading as the automated, computer-based execution of equity orders* via *direct market-access channels, usually with the goal of meeting a particular benchmark."*

Definitions in academic literature dealing with HFT or HFT strategies are mostly based on a "means-to-an-end"-perspective where the means is a large number of transactions in a short period of time and the end are profits in proprietary trading. In this context, Jarnecic and Snape [4] state that: "*HFT is the use of high-speed computer algorithms to automatically generate and execute trading decisions for the specific purpose of making returns on proprietary capital.*" Brogaard [5] puts the emphasize on the means: "*HFT is a type of investment strategy whereby stocks are rapidly bought and sold by a computer algorithm and held for a very short period […] and try to close the trading day in a neutral position"* and Jovanovic and Menkfeld [6] on the end: "*Electronic limit order markets enable agents to automate trading decisions. Computer algorithms are used to either minimize transaction cost when trading into position ('working' an order through time and across markets or to simply profit from buying and selling securities as a middleman). This latter type is the focus of our study and is often referred to as high frequency trading (HFT).*"

Based on a systematic investigation of existing definitions of AT and HFT, Gomber et al. [7] identify the main characteristics that (1) are common to both AT and HFT, (2) are specific for HFT only, and (3) are specific characteristics for the part of AT that is not associated to HFT (see Fig. 13.3).

"Non-HFT" AT is mostly referred to the intelligent working of orders in agent trading relative to a predefined benchmark with holding periods of days, weeks, or

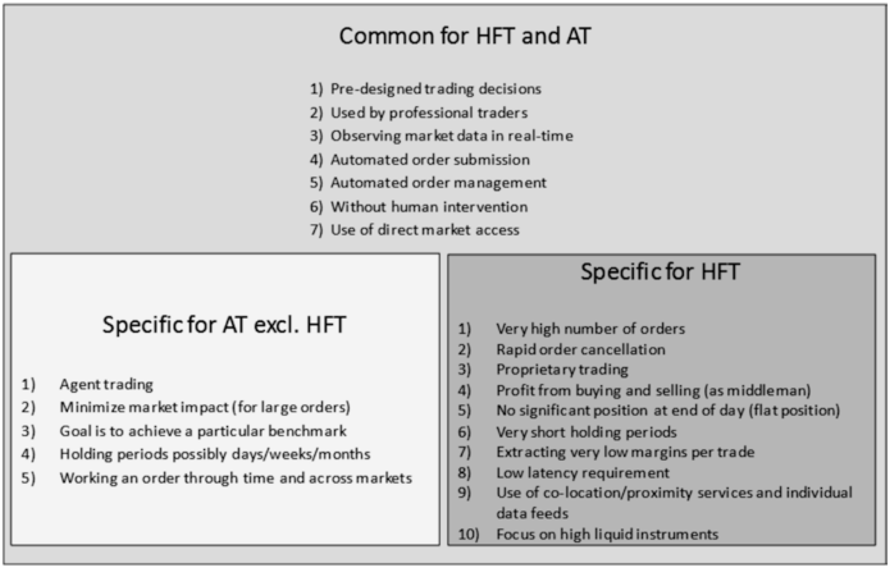


Fig. 13.3 Characteristics of AT and HFT. *Source*: Gomber et al. [7]

even months. In contrast, HFT-based trading strategies update their orders very quickly, avoid over-night positions, and try to realize small profits per trade based on the rapid submission and cancellations of orders. HFT requires high speed access to markets, low latencies and mostly uses co-location/proximity services and individual data feeds. While HFT focuses mainly on high liquid instruments (as these instruments generate the required high number of executions and minimize risk in case of the necessity to unwind positions quickly), AT generates value specifically by minimizing market impact in less liquid instrument like mid or small cap stocks.

The definition of AT and/or HFT in academic papers is often driven by the specific market environment, the analyzed dataset or the concrete research purpose/question of the authors. While the academic exercise of defining and delineating AT and HFT has no direct impact on markets, regulatory definitions of AT and HFT are of high relevance as they trigger specific obligations on the respective firms and on trading venues (for details see Sect. 13.4 of this chapter), thereby directly influencing firms' and trading venues' business models and operations as well as market structure and finally market quality.

The SEC up to today refrained from explicitly defining HFT. However, in its 2010 Concept Release, the SEC gives some insights on its viewpoint by identifying characteristics that are attributed to HFT. An important characteristic is the reference to proprietary traders "*acting in a proprietary capacity that engage in strategies that generate a large number of trades on a daily basis.*"[6] Proprietary traders can be organized in a variety of ways, including as a proprietary trading firm, as the proprietary trading desk of a multi service broker dealer, or as a hedge fund.

Furthermore, the Concept Release specifies the following characteristics that often are attributed to HFT:

1. Use of extraordinarily high speed and sophisticated programs for generating, routing, and executing orders
2. Use of co-location services and individual data feeds offered by exchanges and others to minimize network and other latencies
3. Very short time frames for establishing and liquidating positions
4. Submission of numerous orders that are cancelled shortly after submission
5. Ending the trading day in as close to a flat position as possible

While the US regulators are very cautious in using strict definitions, the EU regulators are clearly more prescriptive in this respect. The first member state of the European Union that explicitly regulated AT and HFT was Germany. On May 15, 2013, the German High-frequency Trading Act entered into effect. It defines both AT and HFT (or more precisely it specifies the definition of an high-frequency algorithmic trading technique): AT is defined in Article 33 (1a) of the German Securities Trading Act: "[…] *trading in financial instruments in such a way that a computer algorithm automatically determines individual parameters of orders, unless the*

[6] See [17–20].

system involved is used only for the purpose of routing orders to one or more trading venues or for the confirmation of orders (algorithmic trading). Parameters of orders within the meaning of sentence 1 include, in particular, decisions on whether to initiate the order, on the timing, price or quantity of the order, or on how to manage the order after its submission with limited or no human intervention."[7]

The definition of a high-frequency algorithmic trading technique is defined both for Article 2 (3) of the German Securities Trading Act and for Article 1 (1a) of the Banking Act. There, high-frequency algorithmic trading techniques are characterized

1. By infrastructures that intend to minimize latency
2. By systems that make the decision to initiate, generate, route, or execute an order without human intervention for individual trades or orders
3. By high intra-day message rates in form of orders, quotes, or cancellations[8]

Only if all three of these criteria are cumulatively fulfilled, a high-frequency algorithmic trading technique is deemed to exist. As "infrastructures to minimize latency" and "high intra-day message rates" are subject to interpretation, BaFin further specified these parameters. The concretization of an infrastructure to minimize latency is based on the distance between the trading venue's matching engine and participant's algo server and on the bandwidth of the connection: as of 2014 this is defined as a 10 Gbit per second transaction line used in co-location. High intraday-message rates means 75,000 messages or more on average over the year per trading venue per trading day determined on a rolling basis per trading day based on the previous 12-month period.[9]

While these definitions are already operational in Germany, the European regulators are in preparations for the 2017 planned go live data of the Markets in Financial Instruments Directive (MiFID II). The framework level 1 text of MiFID II is in effect since July 2, 2014. In MiFID II, AT is defined in Article 4 (39) by stating that: "*Algorithmic trading' means trading in financial instruments where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price or quantity of the order or how to manage the order after its submission, with limited or no human intervention, and does not include any system that is only used for the purpose of routing orders to one or more trading venues or for the processing of orders involving no determination of any trading parameters or for the confirmation of orders or the post-trade processing of executed transactions."*[10]

Very similar to the German approach, the European legislators sees HFT to be a sub-set of AT and specifies the term "high-frequency algorithmic trading technique"

[7] See [21].

[8] See [21].

[9] See [21].

[10] See [22].

to be an algorithmic trading technique characterized by: (1) infrastructure intended to minimize network and other types of latencies, including at least one of the following facilities for algorithmic order entry: co-location, proximity hosting, or high-speed direct electronic access, (2) system-determination of order initiation, generation, routing, or execution without human intervention for individual trades or orders, and (3) high message intraday rates which constitute orders, quotes, or cancellations.

As of late 2014, the concretization of those parameters is still subject to debate. The European Securities Markets Authority (ESMA) has to provide technical advice to the European Commission in the Level II work on MiFID II and has to further specify the definition of what should be considered AT as opposed to a high-frequency algorithmic trading technique, i.e., to also concretize the terms "infrastructure intended to minimize network" and "high message intraday rates." In its 2014 consultation paper, ESMA proposes two options in this respect.[11] While option one is very close to the German parameters (10 Gbit lines; two messages per second over the entire trading day determined on a 12-month rolling basis), a second option was put on the table that applies relative parameters by suggesting to consider participants of a trading venue as being HFTs if their median daily lifetime of orders modified or cancelled is below the median daily lifetime of orders modified or cancelled for the entire market. While this second option has not the disadvantage to use predefined absolute parameters that might be outdated in a few month, its relative approach risks to classify non-HFT members as being HFT just because they are faster than others in a market where HFT does not occur at all or not to a significant extend.

13.2.2 Main Strategies and Players

HFT is not a trading strategy in itself but the usage of sophisticated technology that implements traditional trading strategies. Therefore, any discussion that tries to endorse or prohibit HFT as such or to see HFT as a monolithic structure, as often described in the media, is misleading. Only a well-differentiated assessment of the individual trading strategies can serve as a valid basis for any regulatory intervention. Prohibiting or restricting HFT-based trading strategies that contribute to liquidity or to price formation processes would lead to a reduction in market quality. Nevertheless, any strategies or behaviors that impair market integrity, constitute market abuse, or create an uneven and unfair playing field among market participants have to be banned from our markets. In order to support the argumentation above, in the following, the most recognized HFT-based trading strategies as well as the institutions applying those will be briefly described in the following:

[11] See [23].

1. One of the most common HFT-based strategies is to act as *electronic liquidity providers,*[12] i.e., performing the simultaneous submission of buy and sell orders. Electronic liquidity providers have two basic sources of revenues: (a) spread capturing and (b) monetary incentives by trading venues (rebates). Spread capturing closely resembles traditional market making, i.e., profiting from the spread between the (higher) prices at which market participants buy securities and the (lower) ones at which they sell securities. Rebate-driven strategies are based on maker-taker-pricing models mainly applied by alternative trading venues (ECNs or MTFs). In maker-taker-pricing providing liquidity by passive limit orders (maker) is granted a rebate per executed order, while removing liquidity by sending aggressive, immediately executable orders (taker) is charged a fee (higher than the maker rebate in order to assure a profitable business model for the market operator). Especially for electronic liquidity providers, speed is a central tool to manage inventory risk.
2. *Arbitrage-strategies* are built upon the immediate, high-frequent exploitation of small and short-lived discrepancies between prices of at least two financial instruments or among markets. Examples of those strategies include pairs trading, i.e., current deviations from historical price correlations of stock pairs, cross market trading, i.e., the simultaneous purchase and sell of a financial instrument in different markets which has become more relevant due to increased market fragmentation in the USA and in Europe, and cross asset strategies like the purchase of an exchange traded fund and a parallel shorting of the underlying stocks.[13] Opportunities to conduct arbitrage strategies frequently exist only for fractions of a second. HFTs exploit arbitrage-strategies similar to other non-HFT traders; however they leverage low-latency technology to make use of these (potential) profits before anybody else is able to do so. As arbitrage-strategies have to immediately react to short-lived inefficiencies, they are mainly takers of liquidity. HFT arbitrage strategies are an important means to assure a high alignment of prices among trading venues in the increasingly fragmented market environment thereby largely relieving the ultimate investor from the necessity to intensively compare venues before trading.
3. Another category of HFT-based trading strategies is *liquidity detection*. Applying these strategies, HFTs try to unveil the trading motives and therefore hidden liquidity of other market participants (e.g., hidden orders/iceberg orders or remaining child orders of a large parent order) by submitting orders with a small order size (pinging) or by systematically analyzing the trading activity on market data tickers (sniffing the tape) to detect orders being submitted by other execution algorithms. Based on this information they adjust their own trading behavior at the detriment of the other market participants e.g., by using the knowledge of a large order within the order book as a hedge for the own trading position (quote-matching).[14]

---

[12] See [24].

[13] See [25].

[14] See [26].

4. *Trend-strategies* can be subdivided into short-term-momentum/news reader and latency-arbitrage strategies. Applying short-term-momentum strategies, market participants leverage HFT technologies to conduct strategies equivalent to classical day traders that trade aggressively (taking liquidity) and aim at earning profits from market movements/trends. Trading decisions of these types of HFT can be based on pure market trends or the sophisticated exploitation of economic or company-specific news or events. Such news-reader algorithms automatically digest real-time newsfeeds and analyze structured (e.g., corporate disclosures) as well as unstructured documents (e.g., blogs, tweets, or articles) applying sentiment analysis and text-mining techniques. By applying latency-arbitrage-strategies, HFTs use an information advantage concerning the provision of pricing data from different market places—especially in US markets. Latency arbitrageurs leverage direct data feeds and co-located infrastructure to minimize their reaction times. Since actions of these market participants are said to impair the prices at which other traders (e.g., Buy Side execution algorithms) are able to trade, they are often called "predatory"—see also Sect. 13.5 of this chapter.

There is a highly diverse community of market participants leveraging HFT technologies. A multitude of different institutions with different business models use HFT and there are many hybrid forms ranging from broker-dealer operated proprietary trading firms and broker-dealer market making operations to highly specialized HFT boutiques and quantitative hedge funds. Taking an institutional perspective is therefore misleading. Instead a functional perspective that includes all institutions that apply HFT-based trading strategies assures a level playing field, independent of whether HFT is a core or a supplementary technology for a firm to implement its trading strategies.

13.3 Equity Market Structures in the USA and in Europe

Any developments in trading strategies or any changes in technology have to be discussed in the context of the respective market structure. Often in the discussion around HFT, issues are discussed without distinctively taking into consideration the respective competitive, regulatory and market environment. This has lead to a situation in which (specifically in the media) technological issues in trading—whether correctly attributed to HFT or not—in one market structure environment (e.g., the USA) are transferred to another market environment (e.g., in Europe) without reflecting the specific market structure. This often raises unjustified fears and misunderstandings. Therefore in the following, the main properties of the equity market structures in the USA and in Europe are discussed by analyzing the commonalities but also their distinct differences. Thereafter, recent regulatory initiatives on both sides of the Atlantic will be presented with a specific focus on the regulation of AT and HFT. Based on this analysis, the final section of this chapter will sum up and show that there is an increasing trend towards convergence of market structures in the USA and Europe as regulators, market operators and market participants have transformed the insights from past issues into actions for the future that will increase the safety, integrity, and efficiency of our markets.

13.3.1 Commonalities of US and European Equity Market Structures

Although there is a wide range of potential criteria that can be applied to identify commonalities between the US and European equity market structures, four central factors can be listed in this context:

1. One common regulator
2. Competition among trading venues
3. Consolidation of trading venues
4. High market shares of OTC trading in equities

In contrary to Asia, where each country has its own regulatory setup that is largely independent form other countries, European and US equity market structures are shaped by *one common regulator*. The SEC has played this role since decades, whereas in Europe due to the existence of different EU member states, regulatory harmonization of securities markets proved to be far more difficult. The starting point of European securities market harmonization is the "Investment Services Directive (ISD)" established in 1993.[15] The regulatory framework for investment services and securities trading is developed by the European Commission in a joint process with the European Council and the European Parliament. ESMA, as an independent authority with full accountability towards the Parliament, the Council, and the Commission, is responsible for the development of a single rule book in Europe, to assure investor protection, equal conditions of competition in financial services and effective supervision of financial service providers.[16] Although national competent authorities for securities markets still exist, their role has largely changed to be mainly responsible for supervision and enforcement while the fundamental regulatory decisions on market structure are taken on a European level in Brussels and Paris.

The guiding principle of European and US securities legislation and therefore of both equity market structures is *competition among trading venues*. With Regulation ATS and RegNMS ("Regulation National Market System")[17] in the USA and MiFID[18] in Europe both applicable since 2007, regulators on both sides of the

[15] See [27].
[16] See [28].
[17] See [29].
[18] See [30].

Atlantic intended to create a level playing field between the different types of trading venues and a harmonization in the order execution process. This triggered a fundamental change of the landscape with new market entrants and, thereafter, a wave of consolidation in both regions:

Regulation ATS adopted by the SEC in 1998[19] required Electronic Communication Networks (ECNs)—at that time entering the US market as successful competitors to the incumbent exchanges—to register as broker-dealers or as self-regulated securities exchanges. ECNs have been successful in the USA as they introduced new trading concepts, e.g., the concept of order-driven markets with the open order book approach. This also triggered the incumbents specifically Nasdaq and the NYSE to switch from pure market making or specialist models to hybrid market models integrating an electronic limit order book and liquidity providers (Nasdaq with the rollout of SuperMontage in 2002 and NYSE with its Hybrid Market Initiative in 2005). During the Internet bubble around 2000, multiple ECNs were competing for market share. Given that individual operators struggled to gain market share in a crowded marketplace, a significant *consolidation* among ECNs (e.g., Archipelago acquired REDIbook in 2001; Instinet and Island ECN merged to INET ATS in 2002) and after that a second wave of consolidation among ECNs and exchanges (e.g., Brut LLC and INET ATS were acquired by Nasdaq in 2004 and 2005, respectively; NYSE and Archipelago merged in 2005) took place. As of today, the US market is still highly fragmented with no single venue having a market share of more than 20 % and *OTC trading* representing around 35 % of total equity trading volume.[20]

RegNMS that came into effect in 2007 defines the regulatory framework for the US market structure until today. Key drivers for RegNMS were the issues with the “National Market Systems” (NMS) introduced in 1975 with its key component the “Intermarket Trading Systems” (ITS). ITS linked all US markets trading exchange-listed securities (excluding Nasdaq securities). The concept of best execution in ITS forced markets to route their orders for exchange-listed securities to the market which offered the best price with receiving markets having up to 30 s to respond (“Trade-Through Provision” of the ITS Plan). A Trade-Through occurs when an order is executed on one market despite a better available price on another market. The Trade-Through rule specifically protected markets like the NYSE that provided the most competitive quotes[21] in its listed securities creating the National Best Bid and Offer (NBBO) before the introduction of RegNMS. As these Trade-Through rules were heavily criticized especially by institutional investors that claimed that they don’t reflect the status of technological progress, RegNMS updated the existing rules in order to level the

[19] See [31].

[20] See [32].

[21] See [33].

competitive playing field and by extending the scope of the regulation to Nasdaq stocks. RegNMS classified trading venues in “fast markets” and “slow markets.” The regulation for Trade-Throughs was significantly changed by a new “Order Protection Rule.”[22] Thereby, only orders on fast markets are protected against trade-throughs, i.e., an order can be executed on a “fast market” despite of a better price being available on a “slow market” privileging “fast markets” that enable for automated and immediate execution. “Slow markets” facing a risk of significant loss of order flow modernized their trading systems or merged with markets providing electronic matching facilities (e.g., the merger of NYSE and Archipelago).

In Europe, MiFID came into effect in April 2004 and had to be applied by investment firms and Regulated Markets (RM) since November 2007. A central goal of MiFID is to enable investors to trade securities at maximum efficiency and at minimum cost. This shall be achieved by increasing transparency and accessibility of markets, investor protection, market integrity, harmonized European regulation, and a level playing field among different types of trading venues to assure competition and to foster innovation. MiFID classified trading venues into “Regulated Markets (RM),” “Multilateral Trading Facilities (MTF)” or “Systematic Internalisers (SI).” While RM were already defined in the ISD of 1993 and reflect the incumbent exchanges’ trading setups, MTFs and SIs represent new categories in European securities legislation. An MTF is an alternative trading system bringing together multiple buying and trading interests largely analogous to the ECNs in the USA. A SI is an investment firm that executes customer orders outside a RM or an MTF against the own account.[23]

Before the application of MiFID, a lot of European member states, e.g., Italy, Spain, or France, had concentration rules or required customers to explicitly agree to an execution outside an exchange (“default rule,” e.g., in Germany). This resulted in heavy criticism on exchanges and the accusation of monopolistic or quasi-monopolistic positions that need to be resolved by competition. *Competition among trading venues* triggered by MiFID lead to a highly fragmented European equity market where, as of 2014, the new competitors gained market shares between 20 and 40 % of total lit trading in the main European indices.[24] Although MiFID changed the competitive landscape dramatically, *OTC trading* still represents a high and stable market share of around 40 % of total European equities trading.[25]

While ECNS introduced a new market model in the USA with the electronic clob, the European competition was mainly driven by fee structures (e.g., maker-taker models)

---

[22] See [29].

[23] See [30].

[24] Systematic Internalisiers up to 2014 achieved no relevant market shares (below 3 %; see [34]) and only 12 investment firms are registered as SI [35].

[25] See [32].

and fee levels. European exchanges already had open limit order books and hybrid market models in places before MTFs entered the market. These limit order books are based on the concept of price-time-priority that assures fair and consistent handling of all orders. It is important to point out that price-time-priority assures that time and speed do not matter for limit orders already sitting in the order book as their consistent execution sequence is handled by the trading venues' matching algorithms. However, competition on speed is relevant for market orders or new incoming limit orders.

Multiple industry (e.g., [8]) and academic studies (e.g., Gomber et al. [9], Gresse [10], Hengelbrock and Theissen [11], or Riordan et al. [12]) showed empirically that MiFID not only reduced explicit transaction costs at trading venues and post-trade infrastructures but also achieved increased market liquidity based on a reduction in implicit transaction costs.

Similar to the USA, a multitude of new MTFs entered the market and a *consolidation* among MTFs (e.g., BATS Europe and Chi-X Europe merged in 2011 and were granted status of a Recognized Investment Exchange, i.e., a RM, in May, 2013) and between exchanges and MTFs took place (e.g., the acquisition of the Turquoise MTF by the London Stock Exchange in 2010).

13.3.2 Main Differences Between US and European Equity Market Structures

There are some key differences between the US and the European equity market structures that will be described in this chapter. They include:

1. The consolidation of pre-trade data and the concept of a market wide best quotation
2. The concept of best execution
3. The mechanisms to handle market stress
4. The existence of a home/reference market
5. The clearing and settlement infrastructure

In the USA, RegNMS codified the concept of a national best bid best offer (NBBO). To calculate the NBBO, marketplaces are obliged to distribute their best bid and best offer for securities they are listing to a securities information processor (SIP). The processor *consolidates pre-trade data* by aggregating the quotes coming from marketplaces and ascertains the nationwide best bid and offer in a given security as the NBBO. This *concept of a market wide best quotation* has been implemented to enable market participants to trade on the best available prices in the USA in order to preserve fair competition among brokers and dealers as well as their availability to the market participants.[26] However, there

[26] See [17–20].

are significant issues with the SIP and its operations.[27] In Europe, no such mandated "European Best Bid Offer" is in existence and while MiFID II tries to ensure the development of a pan-European consolidated tape for post-trade transparency, no such pre-trade consolidation is in existence nor is it planned.

Linked to the concept of a national best quote, the US market structure has a fundamentally different *concept of best execution*: The Trade-Through rule (Rule 611 RegNMS) was implemented to guarantee that trades are always executed at the best available price. This rule bars marketplaces from trading at prices that are worse for their customers than the NBBO.[28] If a marketplace is not able to match an incoming order at the NBBO or a better price internally, it is forced to route the order to the trading venue that is currently offering the best price or cancel the order instead of routing it away. In Europe, Article 21 of MiFID introduced a principles-based best execution regime fundamentally different to the rules-based US approach. It requires investment firms to execute orders on terms most favorable to the client. Instead of establishing a pan-European best price like the NBBO, MiFID requires investment firms to: "[…] *take all reasonable steps to obtain, when executing orders, the best possible result for their clients taking into account price, costs, speed, likelihood of execution and settlement, size, nature or any other consideration relevant to the execution of the order*."[29] While best execution is "outsourced" to the market venues in the USA, in Europe best execution is an obligation of investment firms towards their customers and any RM or MTF can execute an order it receives independent from the current bids and offers at other trading venues. Instead of a requirement to determine the best execution venue on an order-by-order basis, investment firms in Europe are required to implement best execution as a process, i.e., to establish an order execution system which routes orders to execution venues that have proven to provide best execution on a consistent basis. As a result, investment firms have to specify an order execution policy which provides information to their customers on the venues they use to execute client orders and to explain the factors leading to this choice. However, the European legislation does not oblige investment firms to connect to every trading venue at any costs if connecting costs are disproportionate. Investment firms have to update their order execution policy

[27] The Securities Industry and Financial Markets Association (SIFMA)—the member-driven trade association—representing a collective voice of market participants reveals that the system operating the SIP is outdated and "suffers from a lack of transparency and competition, questions of underfunding, and insulated governance" [36]. This statement followed after the event of August, 2013 when the SIP operator for Nasdaq listed securities experienced a significant system overload facing 26,000 quote updates per available port per second. The internal error in the SIP software code lead to delay of system output messages. In order to prevent information asymmetry, Nasdaq OMX decided to halt trading for 3 h explicitly denying the fault of HFTs [37]. Also SEC head, Mary Jo White, addressed the problem of consolidated data latency which turns out to be inferior to the latency of direct data feeds [38]. Moreover, she admits that the standards of SIPs robustness and resiliency shall be improved.

[28] See [29].

[29] See [30].

on a regular basis and are required to prove on request that they executed a client's orders in accordance with their respective order execution policy.[30]

In the USA, the discussion around *mechanisms to handle market stress* became of high relevance after the 1987 stock market crash. In order to avoid similar events in the future, NYSE introduced circuit breakers in 1988 to enable the market to review the situation in case of market stress and to be able to pause and reassess the information basis. This shall reduce volatility and avoid self-reinforcing panic. The US circuit breakers were specified based on the movement of the Dow Jones Industrial Average (DJIA) and halted the whole market in case of a significant decline of the index. The duration of the halt depended on the extent of the decline and on the time in the trading day. The respective thresholds and length of the halts were set as follows:

(a) 10 % decline before 2 pm: 1 h halt, between 2 and 2.30 pm: 30 min halt, after 2.30 pm: no halt
(b) 20 % decline before 2 pm: 2 h halt, between 2 and 2.30 pm: 1 h halt, after 2.30 pm: market closure for the rest of the day
(c) 30 % decline: closure for the rest of the day

In Europe, the large exchanges introduced so-called volatility interruptions to prevent extreme price movements in the mid-1990s. Volatility interruptions are triggered if the potential next price is outside predefined ranges (static and dynamic) relative to a reference price that is either based on the last price in continuous trading or the last auction price. In terms of the trading process, a volatility interruption is a switch from continuous trading to a call auction or an extension of the call phase of an auction if the potential extreme price movement occurs within the auction. There are five key characteristics that distinguish European volatility interruptions from US circuit breakers:

(a) Volatility interruptions are parameterized and triggered based on the price movements of each individual instrument rather than on the whole index
(b) Volatility interruptions only affect the respective instrument and not the whole market
(c) The price ranges of volatility interruptions are symmetric, while circuit breakers only react to a market decline
(d) As a volatility interruption is a switch to a call auction, the market is not completely stopped like in the case of a circuit breaker; instead, within the call phase of the volatility interruption, indicative auction prices and volumes are distributed in order to support the price discovery process at the end of the volatility interruption
(e) While circuit breakers halt the market for a relevant time (see above) or even stops trading for the day, volatility interruption in Europe only last a few minutes

[30] See [30].

Thereby, the concept of volatility interruptions enables not only for price discovery during the whole trading day (either in auctions or in continuous trading), but also facilitates the pricing of derivates on single stocks, index calculations, and the pricing of derivates on equity indices.

Although Europe has experienced significant fragmentation since 2007, in nearly all equities the market where the initial listing to place, still serves as the *home/reference market* with market shares mostly above 60 % for individual instruments (among all lit venues).[31]

Finally, the *clearing and settlement infrastructure* in Europe is highly fragmented as post-trading processes in Europe have their origins in a patchwork of national systems. In the USA, the mechanisms for clearing and settlement are concentrated at the DTCC. In Europe, domestic settlement systems are efficient within the national boundaries, but the infrastructure, rules, and systems for clearing and settlement differ and result in high costs for settlement of cross-border transactions and inefficient processes because of various barriers.[32] Although various initiatives like Target-2-Securities[33] or the CSD-Regulation try to harmonize European clearing and settlement and to bring down costs, the existence of multiple local legal jurisdictions, tax laws, and currencies will remain a comparative disadvantage to the US market with one regulation and one tax law.

13.4 Recent Markets Structure and Regulatory Initiatives Concerning HFT

Before 2010, HFT was mostly discussed among market experts and the public did not take notice of the evolution of equity markets. However, the May 2010 US flash crash changed the discussion significantly and triggered regulatory initiatives both in the USA and in Europe. Further issues in the US equity markets heated the debate like the fail of the BATS IPO on March 23, 2012, the technical problems at the Facebook IPO on May 18, 2012, and specifically the Knight Order Flood on August 1, 2012, when due to "technical issues" Knight Capital incorrectly sent orders for 150 symbols to the NYSE and an unchecked Knight Algo caused a $440 million pre-tax loss leading to the takeover of Knight by Getco.

May 6, 2010, represents one of the most devastating plunges in recent equities market history. During less than half an hour, the DJIA and many individual equity securities as well as exchange traded funds experienced a sharp drop followed by an immediate recovery of a significant part of these losses. Over 20,000 trades were executed at prices lower than 60 % than the pre-crash level. Buy-side liquidity in E-Mini contracts decreased by 55 % from $6 billion to $2.65. According to the SEC and CFTC, the main

---

[31] See [32].

[32] See [39] or [40].

[33] See [41].

trigger of the event was a trade initiated by a large institutional trader that was willing to sell 75,000 E-Mini contracts which total value was approximated at $4.1 billion. The respective execution algorithm accounted only for the market participation rate which implied that neither time nor price constraints of an order execution were imposed. As a result, the trade was executed extremely rapidly in about 20 min.

The SEC and CFTC report claims that HFTs immediately reacted on this event by accumulating large long positions over a short period of time. This resulted in aggressive selling that amplified the market drop. In only 14 s, HFTs traded over 27,000 contracts that constitute 49 % of trading volume.[34] In individual securities, HFTs supported the rush and were acting as net sellers. Liquidity providers had to widen the spreads in response to the falling market and many investments firms decided to withdraw the quotes and not to participate. At the end of the day, the DJIA closed with a loss of 3.20 %. In the aftermath of this crash, market participants were fast to accuse HFT for it. However, while their behavior may have contributed to the crash, they are not responsible for it.[35] Rather, specific characteristics within the US market structure that allowed for rapid interaction between trading algorithms without appropriate/unified loss control were the main drivers of the crash.

13.4.1 Market Structure and Regulatory Initiatives Relating to HFT in the USA

The implementation of (1) a new circuit breaker regime may be seen as the most significant market structural change that the USA triggered after the flash crash. Further market structure adjustments/current discussions to adjust regulation relate to the handling of (2) erroneous trades, (3) stub quotes, and (4) the consolidated audit trail system.

Before the flash crash, the USA applied only market-wide circuit breaker based on the narrow DJIA Index. This mechanism was triggered only once in 1997. During the flash crash, the existing circuit breaker regime failed to halt the market and allowed for a rapid decline of individual securities and the overall market. As a result, in April 2011 US regulators introduced a pilot "limit up—limit down" (LULD) program which implements automated circuit breakers for individual securities and updates the rules for the market-wide circuit breaker. The proposed *new circuit breaker regime* establishes price bands around the average quoted levels during the preceding 5-min interval and prevents execution outside of the bands. The price bands vary from 5 % for liquid securities to 10 % for other listed securities and are doubled during the opening and closing periods.[36] The rules concerning the market-wide circuit breaker were also updated: In June 2012, the SEC introduced a new requirement lowering the decline thresholds triggering the market halt to 7, 13, and 20 % from the last closing price rather than 10, 20, and 30 % and simultaneously

[34] See [42].

[35] See [43].

[36] See [44].

shortened the halt times to 15 min rather than closing the market. Moreover, the price references are now based on the broader S&P 500 Index. The newly established rules are more sensitive with respect to overall and individual volatility and take up important positive aspects of the European volatility interruptions. Market participants such as exchanges, trading venues, and dealers that execute trades internally shall have in place mechanisms that follow the LULD rules preserving market integrity and reliability. In April 2014, the SEC decided to extend the pilot program to February 2015.

The annulment of *erroneous trades* also emerged as an issue in the flash crash. In order to provide US markets with a clearer definition of when and under which conditions trades may be broken in the future, the SEC proposed more precise rules for clearly erroneous trades.[37] For stocks priced under $25, trades are defined to be erroneous if trades are at least 10 % away from the circuit breaker trigger price. The requirements for $25–$50 and $50+ securities are 5 % and 3 %, respectively. In case the circuit breaker is not applicable, the quantity traded is considered to be another decisive factor for breaking trades. During the flash crash, orders that were executed at prices 60 % away from the pre-shock levels were identified as erroneous. In January 2013, the SEC addressed this problem of late erroneous order recognition by introduction of a "Multi-Day Event." This mechanism treats a sequence of trades as one event if they "*were affected based on the same fundamentally incorrect or grossly misinterpreted issuance information resulting in a severe valuation error.*"[38] Moreover, if the valuation error is not corrected before the trading halt, all these transactions are declared to be erroneous and are supposed to be nullified. Using this extension to the existing rule, the SEC states that the erroneous orders should not have executed and the market participants shall realize that these kinds of trades will be invalidated which shall lead to more transparent and equitable principles of trade as well as confidence in executions when trading in volatile situations.

During the flash crash, many liquidity providers have started to quote securities far away from the pre-decline levels (e.g., at $0.01 or $100.00). Such quotes are called *stub quotes*. Market makers use them to comply with existing quotation obligations; however, they do not have any intention for these quotes to be executed. Yet, the sudden loss of liquidity due to registered and non-registered market makers withdrawing their quotes lead to the execution of stub quotes which were nullified afterwards. As a result, the SEC has declared a ban on such quotes, forcing market makers in exchange-listed equities to quote within predefined bands around the current NBBO.[39] The maximum distance of the band from the NBBO is declared to be 30 %.

The *consolidated audit trail system* intends to increase the monitoring abilities of US regulators. Aimed at giving regulators the ability to monitor for abuse and analyze atypical events across the fragmented US markets, this system is proposed to be built around one new central database. This database would enable regulators to access detailed information about the lifecycle of each order and quote starting

[37] See [17].

[38] See [38].

[39] See [18].

"*from receipt or origination, through the modification, cancellation, routing and execution of an order*"[40] by 8 am the following trading day. Much of this information would have to be reported in near real time and would need to include "*the ultimate customer who generated the order*,"[41] enabling regulators to conduct investigations faster and more efficiently. For this reason all broker-dealers, national securities exchanges, and even account holders shall be uniquely identified by a code and have in place synchronized business clocks to deliver any reportable event with the respective time stamp.[42] Thus, the consolidated audit trail system facilitates the accumulation/collection of all the possible information regarding financial markets, participants, and their activity across exchange-listed equities and equity options. The data reporting is an obligation that arises from the willingness of the regulators to monitor and analyze the market conditions, players' behavior and assess the effect on market structure of new rules imposed by the SEC.

The aforementioned bans on stub quotes as well as the introduction of a policy to break clearly erroneous trades combined with the improvement of the circuit breaker system pose major steps towards a more resilient, transparent, and stable market system in the USA.

13.4.2 Market Structure and Regulatory Initiatives Relating to HFT in Europe

Although Europe did not experience events like the flash crash or the bankruptcy of a significant market participant (Knight) due to insufficiently tested algorithms, politicians, regulators as well as individual market participants and trading venues took the US incidents as a basis to establish mechanisms that aim at preventing failures due to automated trading algorithms. The most significant ones are the ESMA Guidelines on systems and controls in an automated trading environment for trading platforms, investment firms, and competent authorities that are in place since 2012[43] and MiFID II that came into effect in 2014 and is planned to be applied by market participants as of January 3, 2017. The details of the German HFT law that came into force in May, 2013, will not be discussed in the following as most aspects of the German HFT law will be covered by MiFID II on a pan-European basis.

The ESMA Guidelines include obligations for trading platforms, investment firms, and competent authorities to maintain orderly and functioning markets. The rules enforce a number of organizational requirements to support fair and orderly trading while preventing market abuse, especially market manipulation. In particular, HFTs need to assure that their processes are compliant to regulatory obligations

[40] See [19].

[41] See [19].

[42] See [45].

[43] See [46].

and that their trading systems are resilient, correctly sized, tested, and monitored. The Guidelines demand HFTs to store their trading records and algorithms for 5 years and to employ skilled staff. HFTs need to make sure that automated trading activities comply with regulatory requirements and that the risks resulting from these activities are actively managed. Policies and procedures should be put in place to prevent market abuse. Also, the Guidelines significantly expand the responsibility of trading platforms to monitor and understand the behavior of trading algorithms of market participants. While the ESMA guidelines do not constitute European legislation, they facilitate the consistent interpretation and implementation and specify how the law in the European Union should be applied. However, it is up to each national competent authority to specify if and within what timescale they apply these Guidelines.

In contrast, MiFID II and MiFIR[44] constitute European legislation that has to be either applied directly (in the case of a regulation like MiFIR) or after a transposition into national law by EU member states (in the case of a directive like MiFID II where national implementation has to take place until mid-2016). There are three key pillars within MiFID II that will regulate Algorithmic and High Frequency Trading in the future:

1. Bringing HFTs under the scope of MiFID
2. New specific organizational requirements for investment firms that engage in AT
3. New specific organizational requirements for trading venues. Some of these aspects take up the ideas of the ESMA Guidelines; however, MiFID II introduces additional and more far-reaching requirements

While MiFID[45] exempted entities only dealing on own account (unless they are market makers or provide systems to deal with third parties) from the scope of the directive, MiFID II Article 4 brings into its *scope* all members or participants of a RM or MTF, entities with direct electronic access to a trading venue and entities applying a high-frequency algorithmic trading technique (in the following referred to as HFT; see the discussion on the definitions of HFT above). Therefore, all authorization as well as operating conditions relevant for investment firms also applies to HFTs which will extend regulatory oversight while in parallel increasing operational and regulatory costs for those firms.

Most of the *new specific organizational requirements for investment firms* not only apply to HFT but to all firms applying AT. These organizational requirements (Article 17) include the establishment of systems and risk controls for the operation of the AT systems of investment firms (e.g., concerning capacity, thresholds and limits, business continuity arrangements, and the prevention of market abuse) and the provision of information to competent authorities (e.g., on trading parameters, limits, testing procedures, and record keeping). Furthermore, firms applying algorithmic market making strategies shall provide continuous liquidity on a trading venue based on a market making agreement with the respective trading venue.

[44] See [47].

[45] See [30].

The market making arrangements referred to in Article 17 mostly apply to liquid instruments. Therefore, it is not the intention to increase liquidity in less liquid instruments but to prevent incidents like the flash crash in Europe[46] and to assure that liquidity is not only accessible in "sunshine markets" but also present in times of market stress.

The third pillar refers to *new specific organizational requirements for trading venues*. They are codified in Articles 48–50 for RM. Article 18 requires operators of an MTF or OTF[47] to also comply. Trading venues have to ensure resiliency of their systems, i.e., that systems have sufficient capacities, ensure orderly trading in situations of market stress, are fully tested, and are subjective to business continuity arrangements. Furthermore, trading venues shall assure effective handling of errors by applying volume and price thresholds, rejecting erroneous orders, and be able to cancel or correct transactions. While the implementation of mechanisms to handle volatility was already realized by most European venues on their own initiative since the mid-1990s, MiFID II now explicitly requires venues to have in place mechanisms for trading halts in case of significant price movements. In contrary to the existing mechanisms that are only focusing on volatility within the respective market, MiFID II requires halts also in case of significant price movement on related markets which introduce the need for information and co-ordination among venues. Trading venues shall mirror the market making requirements on investment firms mentioned above by setting up written agreements with those firms including the respective obligations and incentives as well as schemes that ensure that sufficient firms participate in those agreements. Venues have to ensure that members carry out appropriate testing of their algorithms, limit order to transactions ratios, and adopt and enforce minimum tick sizes or tick size regimes that will be harmonized across Europe. In order to enable competent authorities to evaluate algorithmic strategies and to prevent market abuse or risks to orderly functioning markets, trading venues have to require flagging of algorithmic orders by their members.

13.5 Conclusion: Regulatory and Market Structure Convergence

Market fragmentation and technological innovations have been key drivers of the increasing importance of AT and HFT in the last years on both sides of the Atlantic. The fragmentation of liquidity triggered by MiFID and RegNMS has lead to significant structural changes and enabled for new trading strategies that

[46] See [48].

[47] An Organised Trading Facility OTF is a new multilateral trading venue category introduced by MiFID II for non-equity instruments that tries to cover all organized forms of trading that are not organised as a RM or MTF.

were not available or profitable pre fragmentation. The need for HFTs to use speed as the central tools for risk management in electronic liquidity provision and the fact that arbitrage opportunities among products or markets are only exploitable for members with high speed access lead to an arms race to trade at lowest latencies.

US and European markets have been affected differently by incidents related to trading technology due to different market structures. In the USA, events like the flash crash or the Knight bankruptcy triggered immediate reactions targeting to avoid similar situations in the future. There are some specific properties of the US market structure, e.g., the interconnectedness of venues via the Trade-Through rule and the best execution requirements. These specificities increase the predictability of order routing in the USA and triggered strategies like latency arbitrage or market data arbitrage that try to exploit this predictability. Furthermore, the interconnectedness of US markets increases the likelihood of a spillover of market stress as could be witnessed in the flash crash.

In Europe, incidents related to trading technology did not have these significant impacts. Although fragmentation and competition is a reality also in European equity trading, the predominant market in terms of market share and liquidity is still the primary listing venue. Furthermore, a different best execution regime and decentralized order handling does not enable for the strategies named above.

While in the USA, a lot of the new regulations around trading technologies directly resulted from the experiences of the flash crash, European market participants, venue operators, and regulators tried to infer the necessary actions form the US experiences and thereby to increase market stability and integrity.

The experiences and the actions taken lead to a tendency towards an increasing convergence of market structures and regulations in both setups. Similar approaches and the identification and implementation of best practices can be identified in various fields: e.g., the handling of extreme market stress on an instrument-by-instrument basis via LULD rules and volatility interruptions respectively, dealer registration and licensing, requirements on risk management and testing by market participants and trading venues as well as the debate around market making obligations for HFT firms.[48]

While a lot of initial discussions questioned the economic justification of these trading technologies in general and even proponents of a complete prohibition raised their voice, meanwhile the debate has become more differentiated and balanced [49]. Various studies provided by academics and market operators enable for evidence-based regulation and a clearer picture on the impact of new trading technologies on market quality. The actions taken in the last years show that regulators and market operators on both sides of the Atlantic are intending to preserve the benefits of HFT while trying to mitigating the risks as far as possible.

[48] See [38].

References

1. T. Hendershott, C. M. Jones and A. J. Menkveld, "Does algorithmic trading improve liquidity?", Journal of Finance 66, no. 1, 2011, pp. 1–33.
2. A. Chaboud, B. Chiquoine, E. Hjalmarsson and C. Vega, "Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market", Journal of Finance 69, no. 5, 2014, pp. 2045–2084.
3. I. Domowitz and H. Yegerman, "The Cost of Algorithmic Trading: A First Look at Comparative Performance", Journal of Trading 1, no. 1, 2006, pp. 33–42.
4. E. Jarnecic and M. Snape, "An analysis of trades by high frequency participants on the London Stock Exchange", 17th Annual Conference of the Multinational Finance Society MFS 2010, 2010.
5. J. Brogaard, "High Frequency Trading and its Impact on Market Quality", 5th Annual Conference on Empirical Legal Studies, 2010.
6. B. Jovanovic and A. J. Menkveld, 2010. Working Paper: "Middlemen in Limit-Order Markets". Available at: http://ssrn.com/abstract=1624329.
7. P. Gomber, B. Arndt, M. Lutat and T. Uhle, 2011. Working Paper: "High-Frequency Trading". Available at: http://ssrn.com/abstract=1858626.
8. Oxera, 2011: "Monitoring prices, costs and volumes of trading and post-trading services". Available at: http://www.oxera.com/Oxera/media/Oxera/downloads/reports/Oxera-report-on-trading-and-post-trading-May-2011.pdf?ext=.pdf.
9. P. Gomber, M. Gsell and M. Lutat, "Competition among electronic markets and market quality", 14th Conference of the Swiss Society for Financial Market Research (SGF), 2011.
10. C. Gresse, 2012. Working Paper: "Effects of Lit and Dark Trading Venue Competition on Liquidity: The MiFID Experience". Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1918473.
11. J. Hengelbrock and E. Theissen, 2009. Working Paper: "Fourteen at One Blow: The Market Entry of Turquoise". Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1570646.
12. R. Riordan, A. Storkenmaier and M. Wagener, 2011. Working Paper: "Do Multilateral Trading Facilities Contribute to Market Quality?". Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1852769.
13. P. Gomber and K. Zimmermann: "Algorithmic Trading", in: S.-H. Chen and M. Kaboudan, OUP Handbook on Computational Economics and Finance, Oxford University Press, 2014.
14. M. Pagano and A. Roell, "Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading", Journal of Finance 51, no. 2, 1996, pp. 579–611.
15. P. K. Jain, "Financial Market Design and the Equity Premium: Electronic versus Floor Trading", Journal of Finance 60, no. 6, 2005, pp. 2955–2985.
16. B. Johnson, Algorithmic Trading and DMA: An introduction to direct access trading strategies, 4Myeloma Press, 2010.
17. SEC—Securities and Exchange Commission, 2010: "Concept Release on Equity Market Structure". Available at: http://www.sec.gov/rules/concept/2010/34-61358fr.pdf.
18. SEC—Securities and Exchange Commission, 2010. Press Release: "SEC to Publish for Public Comment Proposed Rules for Clearly Erroneous Trades". Available at: http://www.sec.gov/news/press/2010/2010-104.htm.
19. SEC—Securities and Exchange Commission, 2010 "SEC Approves New Rules Prohibiting Market Maker Stub Quotes". Available at: http://www.sec.gov/news/press/2010/2010-216.htm.
20. SEC—Securities and Exchange Commission, 2010: "Speech by SEC Chairman: Opening Statement at the SEC Open Meeting—Consolidated Audit Trail". Available at: http://www.sec.gov/news/speech/2010/spch052610mls-audit.htm.
21. BaFin—Bundesanstalt für Finanzdienstleistungsaufsicht, 2014: "Translation of the main provisions of the High Frequency Trading Act (Hochfrequenzhandelsgesetz), non-official translation for information purposes only". Available at: http://www.bafin.de/SharedDocs/Aufsichtsrecht/EN/Gesetz/hft_en.html?nn=2821360#doc4708178bodyText24.

22. MiFID II, 2014: "Directive 2014/65/EC of the European Parliament and of the Council". Available at: http://europa.eu.int/eur-lex/pri/en/oj/dat/2004/l_145/l_14520040430en00010044.pdf.
23. ESMA—European Securities and Markets Authority, 2014: "Consultation Paper MiFIDII/MiFIR". Available at: http://www.esma.europa.eu/system/files/2014-549_-_consultation_paper_mifid_ii_-_mifir.pdf.
24. ASIC—Australian Securities and Investments Commission, 2010. Report 215: "Australian equity market structure". Available at: http://www.asic.gov.au/asic/pdflib.nsf/LookupByFileName/rep-215.pdf/$file/rep-215.pdf.
25. I. Aldridge, High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems, John Wiley & Sons, 2010.
26. L. Harris, Trading and exchanges. Market microstructure for practitioners, Oxford University Press, 2003.
27. European Council, 1993. "Council Directive 93/22/EEC on investment services in the securities field". Available at: http://www.esma.europa.eu/system/files/Dir_93_22.PDF.
28. ESMA—European Securities and Markets Authority, 2014: "How ESMA works". Available at: http://www.esma.europa.eu/page/esma-short.
29. SEC—Securities and Exchange Commission, 2005: "Regulation NMS Final Rule: Release No. 34-51808, June 9th". Available at: http://www.sec.gov/rules/final/34-51808.pdf.
30. European Union, 2004. Markets in Financial Instruments Directive released April 21st: "Directive 2004/39/EC of the European Parliament and of the Council". Available at: http://www.esma.europa.eu/system/files/FIMD___official_Journal_version.pdf.
31. SEC—Securities and Exchange Commission, 1998: "Regulation of Exchanges and Alternative Trading Systems: Final Rule Release No. 34-40760 December 8th". Available at: http://www.sec.gov/rules/final/34-40760.txt.
32. Thomson Reuters, 2014: "Monthly Market Share Reports". Available at: http://thomsonreuters.com/monthly-market-share-reports/.
33. NYSE, 2006: "NYSE Announces 2005 Financial Results". Available at: http://www1.nyse.com/press/1138706382586.html.
34. Fidessa, 2014: "Fidessa Fragmentation index". Available at: http://fragmentation.fidessa.com.
35. ESMA—European Securities and Markets Authority, 2014: "MiFID Database". Available at: http://mifiddatabase.esma.europa.eu.
36. SIFMA, 2013: "Letter from Theodore R. Lazo, Managing Direct and Associate General Counsel, SIFMA to Mary Jo White, Chair, Commission dated December 5, 2013". Available at:http://www.sifma.org/comment-letters/2014/sifma-writes-letter-to-nasdaq-utp-plan-operating-committee-on-selection-of-processor-for-the-nasdaq-sip/.
37. Nasdaq OMX, 2013: "NASDAQ OMX Provides Updates on Events of August 22, 2013". Available at: http://ir.nasdaqomx.com/releasedetail.cfm?ReleaseID=787888.
38. SEC—Securities and Exchange Commission, 2014: "Enhancing Our Equity Market Structure, Speech by Chair Mary Jo White, Sandler O'Neill & Partners, L.P. Global Exchange and Brokerage Conference New York, June 5, 2014". Available at: http://www.sec.gov/News/Speech/Detail/Speech/1370542004312#_edn6.
39. Giovannini Group, 2001: "Cross-Border Clearing and Settlement Arrangements in the EU". Available at: http://ec.europa.eu/internal_market/financial-markets/docs/clearing/first_giovannini_report_en.pdf.
40. H. Schmiedel, M. Malkamäki and J. Tarkka, "Economies of scale and technological development in securities depository and settlement systems", Journal of Banking & Finance 30, no. 6, 2006, pp. 1783–1806.
41. European Central Bank, 2007. "TARGET2-Securities—The Blueprint". Available at: https://www.ecb.europa.eu/pub/pdf/other/t2sblueprint0703en.pdf.
42. CFTC & SEC, 2010. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues (September 30): "Commodity and Futures Trading Commission & Securities and Exchange Commission, Findings Regarding the Market Events

of May 6, 2010". Available at: http://www.sec.gov/news/studies/2010/marketevents-report.pdf.
43. A. A. Kirilenko, A. S. Kyle, M. Samadi and T. Tuzun, 2014. Working Paper: "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market". Available at: http://ssrn.com/abstract=1686004.
44. SEC—Securities and Exchange Commission, 2011: "Release 2011-84: SEC Announces Filing of Limit Up-Limit Down Proposal to Address Extraordinary Market Volatility". Available at: http://www.sec.gov/news/press/2011/2011-84.htm.
45. SEC—Securities and Exchange Commission, 2012: "Release 201-134: SEC Approves New Rule Requiring Consolidated Audit Trail to Monitor and Analyze Trading Activity". Available at: http://www.sec.gov/News/PressRelease/Detail/PressRelease/1365171483188.
46. ESMA—European Securities and Markets Authority, 2012. Regulatory Report: "Guidelines on systems and controls in an automated trading environment for trading platforms, investment firms and competent authorities". Available at: http://www.esma.europa.eu/system/files/2011-456_0.pdf.
47. MiFIR, 2014: "Regulation (EU) No 600/2014 of the European Parliament and of the Council". Available at: http://europa.eu.int/eur-lex/pri/en/oj/dat/2004/l_145/l_14520040430en00010044.pdf.
48. ESMA—European Securities and Markets Authority, 2014: "Discussion Paper MiFIDII/MiFIR". Available at: http://www.esma.europa.eu/system/files/2014-548_discussion_paper_mifid-mifir.pdf.
49. SEC—Securities and Exchange Commission, 2014: "Release No. 34-72434". Available at: http://www.sec.gov/rules/sro/bats/2014/34-72434.pdf.

Chapter 14
Global Developments in Equity Trading

Phupinder Gill

14.1 Global Developments in Equity Derivatives

Global equity markets have experienced a dramatic transformation over recent years, spurred by rapidly advancing electronic trading technologies, intense competition, and a challenging regulatory environment. Equity derivatives markets have likewise evolved rapidly. These derivatives include options on individual equities and stock indexes; futures on stock indexes and individual equities; exchange traded funds (ETFs); over-the-counter (OTC) equity swaps; structured derivatives; and some rather unique new products based upon volatility and dividends.

Thus, the marketplace now offers more diverse ways in which to attain and manage equity risk exposures than ever before. This chapter is intended to provide a review of how equity derivatives emerged and have evolved to the current day. Along the way, we will underscore the famous quote by Harry Truman—"the only thing new in the world is the history you do not know."

14.1.1 Form and Underlying Subject

Equity derivatives may be distinguished on the basis of the form they assume, the underlying instrument upon which they are based, or even the regulatory jurisdiction under which they are offered. To illustrate, equity derivatives may be offered in the form of a futures contract, an option, or a security such as an ETF or other imaginative combinations. Today's financial engineers are constantly creating new concepts—by bundling or unbundling existing structures—or creating altogether new forms of derivatives.

P. Gill (✉)
CME Group Inc., 20 South Wacker Drive, Chicago, IL 60606, USA
e-mail: Phupinder.Gill@cmegroup.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_14

As the name implies, “derivatives” are financial instruments that are based upon, or derived from, some preexisting underlying subject or instrument. The fundamental building block of the equity marketplace is found in a share of stock. Thus, many equity derivatives are based directly upon a singular stock issue. But other equity derivatives may be based upon a combination of stocks. The earliest examples of such combinations are found in the mutual fund industry. But the concept of indexation and a passive investment strategy have spurred exciting new developments in this regard.

Finally, derivatives are subject to the regulatory authorities in the jurisdiction in which they operate or are registered. Thus, regulatory considerations have frequently played a significant role in shaping the course of derivative product development.

The process of developing new equity derivatives to serve the needs of the marketplace is marked by frequent false starts and failures. But all initiatives are driven by a common goal—to create more efficacious and efficient investment and risk-management opportunities.

14.1.2 Mutual Funds as the Original Equity Derivatives

While equity mutual funds are not classically considered “derivatives” per se, they might be considered the very first form of equity derivatives. The origin of the mutual fund is probably rooted in the financial upheavals associated with European colonial expansion in the eighteenth century. The British East India Company was chartered in 1600 and grew to account for perhaps half of all world commodity trade. But by 1773, Europe was in the midst of an economic depression and the company’s vast interests in India and the New World had soured. The British Parliament, recognizing that the company had become “too big to fail,” was compelled to step in with a restructuring plan and financial bailout.

It was against that backdrop of economic turmoil that the Dutch merchant Adriaan van Ketwich created the first closed-end investment trust in 1774 for purposes of mutualizing risks from colonial plantation investments. The concept was simple—create a diversified portfolio of equities and sell shares for proportionate interests in that bundle of equities. That philosophy was reflected in the name of the trust—Eendragt Maakt Magt—which means “unity makes strength.” Subsequently King William I of the Netherlands created several closed-end investment companies in 1822. This was followed by the development of other investment trusts in Switzerland in 1849 and Scotland in the 1880s.

The Boston Personal Property Trust was founded in 1893 and represents the first recorded example of a closed-end mutual fund in the USA. This was followed by the Alexander Fund in 1907 which allowed investors to add funds semi-annually and withdraw funds on demand. The Massachusetts Investors’ Trust was introduced in 1924, went public in 1928 and is thought of as the first modern open-ended mutual fund. By 1929, there were over 700 mutual funds in the USA—mostly leveraged closed-end funds with a handful of open-ended funds. Many of these funds did not survive the crash of 1929 but the concept survived and eventually thrived.

Equity mutual funds are now widely embraced as a fundamental investment vehicle. The Investment Company Institute (ICI) estimates that 44 % of all mutual fund assets are devoted to equity investment, representing some $13.6 trillion in value as of the 1st quarter 2014.

14.1.3 Stock Indexes and Passive Investment Strategy

Stock indexes represent an extension or variation on the basic theme of a mutual fund to the extent that they simply represent a combination of individual equities. A large proportion of equity derivatives are based upon stock indexes. Thus, a word is in order regarding the popularity of indexation as a benchmark of performance as well as the basis for passive investment strategies.

Stock indexes trace their origin to journalists Charles Dow and Edward Jones who founded the Wall Street Journal. In 1884, Dow began publishing a stock average comprised of nine railroad and two industrial stocks that was the precursor of what became the Dow Jones Transportation Average. By 1886, Dow created the Dow Jones Industrial Average (DJIA), originally a simple average of 12 industrial issues.

The concept soon gained momentum and by the early twentieth century, indexes proliferated around the globe. Interest in stock averages or indexes was driven by the popularity of Dow Theory, embodied in a collection of articles by Dow suggesting that market trends may be forecast by past movement in the averages, as well as an increasing use of statistical and mathematical techniques in the field of economics.

Indexes were further popularized when a group of economists, led by William Sharpe, espoused the Capital Asset Pricing Model (CAPM) beginning in the early to mid-1960s. The CAPM compares the risks associated with any individual equity with systematic market risks, or beta (β), which are frequently represented by a "benchmark" stock index.

The passive investment strategy, whereby an investor simply buys and holds the constituents of a benchmark stock index to capture "beta" returns, flows from the simplicity and enduring appeal of the CAPM. Passive management suggests that investors need not attempt to pick stocks or frequently reconstitute their equity portfolios to achieve attractive returns. Rather, one may achieve the returns generated by the market as a whole while minimizing portfolio turnover, associated execution costs, and management fees. Sharpe suggests that the net result is that "the average actively managed dollar must underperform the average passively managed dollar, net of costs[1]."

The first index mutual fund was created in 1971 by William Fouse and John McQuown of Wells Fargo Bank. Barclays Global Investors (BGI) explored index related investment products as early as 1971 and continues, under the moniker iShares by BlackRock, to be the dominant ETF sponsor in the industry. Other index-

[1] Sharpe, William, "The Arithmetic of Active Management," The Financial Analysts Journal, Vol. 47, No. 1, January/February 1991.

based mutual funds quickly followed including the venerable Vanguard 500 Index Fund (VFIAX) which is designed to track the performance of the Standard & Poor's 500 (S&P 500), managing some $107 billion in assets as of August 2014.

14.1.4 Individual Equity Options and Stock Index Options

Options are not a new financial concept. Aristotle's Politics (circa 350 B.C.) describes how Thales effectively purchased what amounted to a primitive form of an option to rent olive presses in anticipation of a bumper crop the next season. Options were similarly utilized during the tulip bulb bubble of 1637 in Holland as a means of levering investments. Stock option trading has been conducted in the USA for almost as long as there have been stocks. Traditionally, stock options were traded on an OTC basis through a loose network of broker-dealers known as the Put-Call Dealers Association.

The modern model for option trading did not emerge until 1973 when the Chicago Board of Trade (CBOT) established the first exchange organized exclusively for the trade of options in the form of the Chicago Board Options Exchange (CBOE). CBOE was organized by CBOT as a separately managed entity because of regulatory reasons. Specifically, the Securities and Exchange Commission (SEC) regulates securities and options on securities in the United States. The Commodity Futures Trading Commission (CFTC) generally regulates futures, including stock index futures and options on stock index futures.

Individual Stock Options—The launch of CBOE represented a seminal event in the history of options to the extent that CBOE established the modern model by which options would be traded thenceforth. A call option represents the right to buy the underlying instrument, typically 100 shares of a specified stock, at a fixed strike or exercise price, on or before a fixed expiration date. A put option represents the right to sell the underlying instrument at a fixed strike price on or before the expiration date. The option buyer compensates the seller by paying a negotiated premium. The option buyer then enjoys the right to exercise the option or not while the seller (writer or grantor) is obligated to take the opposite side of the trade upon exercise.

In the past, options terminology and trading practices were quite varied. Options were sometimes known as "privileges" or "indemnities," calls were referred to as "bids," puts were known as "offers." Sometimes option premiums, or "commissions," were fixed while the strike or exercise price was variable and determined relative to the current market price. But the success of the CBOE model ushered us into the modern era of option trading.

The model was refined by 1975 when the SEC mandated that the Options Clearing Corporation (OCC) serve as the clearing house for all US exchange-traded stock option transactions. This created fungibility between options traded on different exchanges, i.e., one could establish an option position on one exchange and liquidate it on another exchange to the extent that they were all cleared through OCC. This resulted in a proliferation of competing stock option exchanges in the USA, with some 12 option exchanges operating as of 2013 alongside many others around the world.

Stock Index Options—Options that are directly settled in cash by reference to a spot stock index—as opposed to an option exercisable for a stock index futures contract—are another important part of the mix. Stock index options commenced trading in 1983 on CBOE. Since then, the concept has spread worldwide.

It should be noted that the USA maintains a rather unusual regulatory dichotomy between security derivatives, regulated by the SEC, and futures derivatives, regulated by the CFTC. Per US law and regulation, options that are directly cash-settled to a "broad-based" stock index are regulated by the SEC. Options that are settled when exercised into a stock index futures position are regulated by the CFTC. Thus, stock index options are traded on SEC-registered securities exchanges in the USA while options on stock index futures are offered on CFTC-registered futures exchanges.

This is a subtle but important jurisdictional distinction that is generally not recognized in other countries. Thus, stock index options are frequently traded on the same exchange that offers stock index futures outside of the USA. (Fig. 14.1).

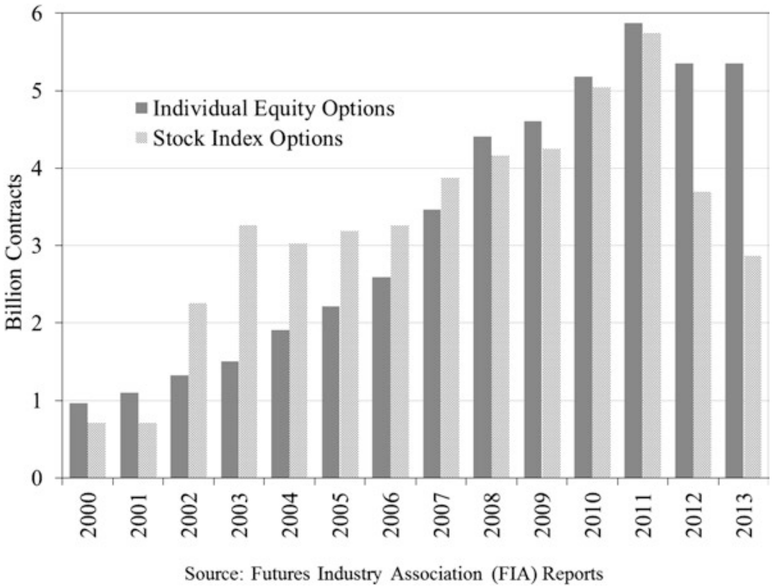


Fig. 14.1 Growth in the global market for stock and stock index options

14.1.5 Stock Index Futures, Options on Futures, and Single Stock Futures

Stock index futures emerged in 1982 as a means to manage the risks associated with diversified stock portfolios. The concept dates back at least to 1970 when the (now defunct) New York Produce Exchange attempted to list futures based on the DJIA, only to be denied by SEC. It was not until 1982 that the requisite regulatory constructs to support stock index futures were in place that the market was able to be introduced in the USA.

Stock index futures combine elements of both the spot equity markets and futures market. As such, the SEC and CFTC had to partition regulatory responsibility, assigning oversight duties for futures on what became known as "broad-based indexes" to the CFTC, reserving responsibility for futures on "narrow-based indexes" and individual equities to the SEC and CFTC jointly. This agreement was known as the Johnson-Shad Accord, by reference to the Chairmen of the two agencies, later to be refined by the Commodity Futures Modernization Act (CFMA) of 2000. Further, the CFTC had to recognize cash-settlement as a legitimate final settlement mechanism. Thus, the regulatory infrastructure required some time to "catch-up" to the imagination of financial engineers and the demands of the marketplace. As such, the regulatory environment had a direct role in shaping new product development.

Stock index futures are financial contracts that are simply marked-to-market on their final settlement date to the spot value of the associated index. They are valued at some nominal monetary amount multiplied by the index value, e.g., an E-mini S&P 500 futures contract is notionally valued at $50 × Index. Thus, if the S&P 500 is at 1900 index points, this implies a notional futures contract value of $95,000 (=$50 × Index). These relatively simple tools are secured with a margin deposit that typically represents 5–15 % of the notional value of the contract.

Stock Index Futures—The Kansas City Board of Trade (KCBT) enjoyed the distinction of listing the very first stock index futures contract in the form of Value Line Composite Average (VLCA) futures. This was followed closely by the introduction of Standard & Poor's 500 (S&P 500) futures on the Chicago Mercantile Exchange (CME) as well as many other contracts around the globe.

Liquid stock index futures are available around the globe including Europe where EUREX lists futures on the pan-European Euro STOXX 50 while NYSE Euronext lists futures on the FTSE 100. Benchmark Asian stock index futures include the Nikkei 225 listed on the Osaka Securities Exchange (OSE); the China Stock Index 300 (CSI-300) on the China Financial Futures Exchange (CFFEX); CNX Nifty futures on the National Stock Exchange of India (NSE); and KOSPI 200 futures and options available on the Korean Exchange (KRX). Latin American stock index products include the Ibovespa listed on BM&F Bovespa in Sao Paulo and the IPC listed on Mexder in Mexico City (Fig. 14.2 and Table 14.1).

Options on Futures—Options on stock index futures were introduced in 1983, nearly simultaneously with securitized stock index options. The two product lines serve similar economic purposes and are traded per very similar practices (Fig. 14.3).

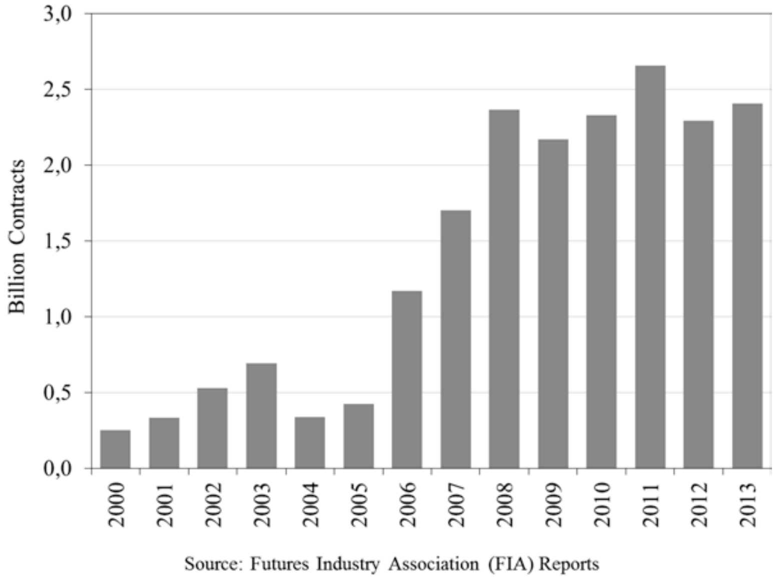


Fig. 14.2 Global volume growth in stock index future

Table 14.1 Top 20 stock index futures contracts ranked by average daily volume (ADV) in 2013

	Contract	Multiplier	Exchange	2013 ADV	ADV (Mil $)	Dec-13 OI	OI (Mil $)
1	E-mini S&P 500 Index	$50	CME	1,794,807	$165,873	2,780,195	$256,940
2	CSI 300 Index	300 CNY	CFFEX	766,748	$88,504	119,534	$13,797
3	Euro Stoxx 50	10 EUR	EUREX	1,065,457	$45,524	2,857,080	$122,074
4	Dax	25 EUR	EUREX	112,769	$37,009	159,290	$52,277
5	Kospi 200	500000 KRW	Korean Exch	198,297	$24,884	139,187	$17,466
6	Nikkei 225 Futures	1000 JPY	SGX	155,110	$23,998	329,407	$50,965
7	Nikkei 225 Futures	1000 JPY	Osaka SE	122,650	$18,976	420,037	$64,988
8	E-mini Nasdaq 100	$20 × Index	CME	235,687	$16,932	419,677	$30,150
9	FTSE 100 Index	10 GBP	LIFFE UK	133,052	$14,868	563,357	$62,952
10	Nikkei 225 Mini	100 JPY	Osaka SE	928,018	$14,358	673,736	$10,424
11	Russell 2000 Mini Index	$ 100 × Index	ICE	114,435	$13,316	296,433	$34,494
12	Mini-sized $5 DJIA	$5 × Index	CBOT	140,670	$11,659	119,639	$9,916

(continued)

Table 14.1 (continued)

	Contract	Multiplier	Exchange	2013 ADV	ADV (Mil $)	Dec-13 OI	OI (Mil $)
13	Topix Index Futures	10000 JPY	Tokyo SE	90,135	$11,148	594,299	$73,502
14	CAC 40	10 EUR	LIFFE Paris	147,666	$8,718	342,977	$20,249
15	S&P 500 Index	$250 × Index	CME	16,494	$7,622	142,221	$65,719
16	H-Shares Index	50 HKD	HKFE	82,822	$5,777	217,646	$15,180
17	Taiex Futures (TX)	200 TWD	Taiwan Fut	90,053	$5,200	66,454	$3,837
18	SPI 200	25 AUD	Sydney Fut	40,633	$4,816	227,654	$26,984
19	MSCI Singapore Index	200 SGD	SGX	14,811	$3,998	38,370	$10,358
20	AEX Stock Index (FTI)	200 EUR	LIFFE Am	35,866	$3,961	59,388	$6,559

ADV average daily volume, which may he expressed in number of contracts traded or as notional value of contracts traded
Source: Futures Industry Association (FIA) Reports

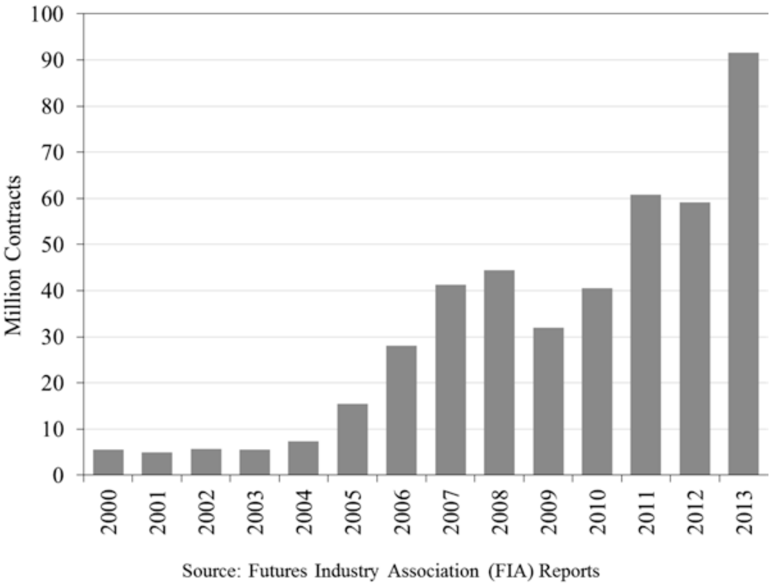


Fig. 14.3 Volume growth in options on stock index futures

Options on futures in securities products, unlike securitized stock index options, are not settled in cash when exercise. Rather, they are settled upon exercise with the delivery of a futures contract that can subsequently be settled in cash at spot index value. But for some period, however brief, the exercised option results in the assignment of a futures contract into the trader's account. As explained above, this is a variation on the theme that is rather unique to the peculiar regulatory environment found in the USA.

Single Stock Futures—Single stock futures are futures that call for the delivery of shares of individual equity securities. While the product has gained new momentum over the past decade or so, the concept is not new. New York stock traders frequently engaged in "time bargains," or what we would refer today as single stock futures contracts, extending back into the eighteenth century. These contracts frequently were satisfied in a cash settlement rather than a physical delivery. In 1812, the New York legislature passed an Act declaring all contracts for the sales of securities void unless the seller at the time of the sale was the actual owner or authorized sales agent for the owner. Time bargains continued to be traded after passage of this legislation although they did not fall under the protection of New York commercial law but rather relied upon New York Stock Exchange (NYSE) rules as the only way to enforce such contracts. By 1858, the New York legislation was repealed although time bargains seemed to wither and die naturally.

Today, time bargains or single stock futures are traded actively on various exchanges including EUREX, Moscow Exchange, National Stock Exchange of India, NYSE LIFFE, Korea Exchange, and OneChicago (Fig. 14.4).

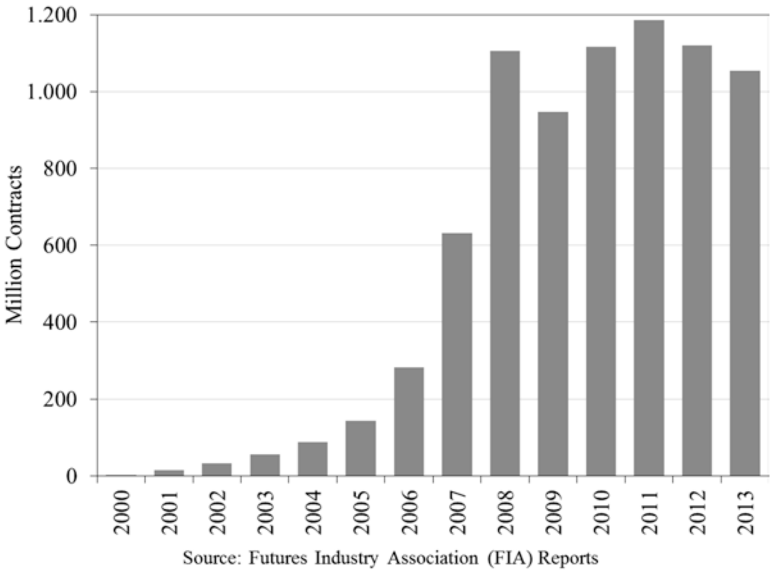


Fig. 14.4 Growth in single stock futures

Note that single stock futures broadly include futures on ETFs to the extent that ETFs may be considered a singular security even though the underlying trusts may hold myriad stock issues.

Actually, single stock futures were traded at various venues across Europe and Asia well in advance of it 2002 emergence in the U.S. Jurisdictional issues between the SEC and CFTC had prevented the introduction of single stock futures in the U.S. until the passage of the CFMA of 2000 which mandated that the agencies develop a joint regulatory framework for such products. This framework was developed and, in 2002, the OneChicago exchange was established, devoted exclusively to these products. This might be characterized as another instance where regulatory infrastructure had to catch-up to the demands of the marketplace.

14.1.6 Exchange Traded Funds (ETFs)

Exchange Traded Funds (ETFs) were originally introduced in 1992 to a lukewarm reception. But by the late 1990s, volumes and assets under management (AUM) started to increase on an accelerated pace. Today, ETFs are widely embraced both domestically and internationally as a mainstream investment vehicle, having won a strong following amongst retail and institutional customers alike. So much so that ETF volume now accounts for about half of all volume on domestic securities exchanges. While ETFs are organized as securities, they may be considered a form of derivatives contract.

While they were first introduced in the context of equities, the concept has been extended to other markets including fixed income items, currencies, and commodities.

ETF History—ETF-like structures trace back to the introduction of Index Participation Shares (IPS), an S&P 500 proxy that traded on AMEX and the Philadelphia Stock Exchange (PHLX). This product was short-lived after courts declared that the IPS product was essentially a futures contract subject to the oversight of the Commodity Futures Trading Commission (CFTC) and not the Securities Exchange Commission (SEC). Once again, the regulatory dichotomy between the CFTC and SEC impacted the course of product development.

A similar product, Toronto Index Participation Shares, started trading on the Toronto Stock Exchange (TSE) in 1990. The shares, which tracked the TSE 35 and later the TSE 100 stocks, proved to be popular. The popularity of these products prompted AMEX to go back to the drawing board and develop a new structure that would fall under SEC jurisdiction and which could be marketed as a security in the United States.

AMEX executive Nathan Most is generally credited with the development of Standard & Poor's Depositary Receipts introduced in January 1993. Known as the SPDR S&P 500 ETF Trust (SPY) or "Spiders," the fund became the largest ETF in the world with $173 billion in net assets as of August 2014.

Other ETFs based on non-US market indexes quickly followed. BGI introduced World Equity Benchmark Shares (WEBS) in 1996, subsequently renamed iShares

MSCI Index Fund Shares. WEBS tracked MSCI country indexes, originally 17, of the funds' index provider, Morgan Stanley Capital International.

While SPDRs were organized as unit investment trusts, WEBS were established as a mutual fund, the first of their kind. In 1998, State Street Global Advisors introduced "Sector Spiders," based on nine industrial sectors drawn from within the S&P 500.

Ever since then, ETFs have proliferated, addressing an increasing array of regions, sectors, commodities, bonds, futures, and other asset classes. The World Federation of Exchanges (WFE) reports that there were 8160 ETFs as of the conclusion of 2013 with an estimated $11.9 trillion in value traded in 2013. That is up from 331 ETFs with an aggregate value traded of $114.5 billion in 2003. The USA remains the epicenter of activity, with 86.5 % of volume transacted on NASDAQ OMX and NYSE in 2013 (Fig. 14.5).

The growth in the ETF industry to its present impressive magnitude is a direct function of intense competition to win marketshare coupled with supporting electronic trading technologies. Certainly this rapid product proliferation could not have been possible in a traditional physical trading environment to the extent that there simply would not have been sufficient floor space or market specialists to support a very wide variety of customized but sometimes less active products.

What Is an ETF?—An index ETF pools the assets of its participants to invest in indexes and meet clearly identified objectives, such as current income or capital appreciation. Index ETFs are passively managed and allow investors to achieve exposure to a portfolio of securities in a single transaction. These products offer an

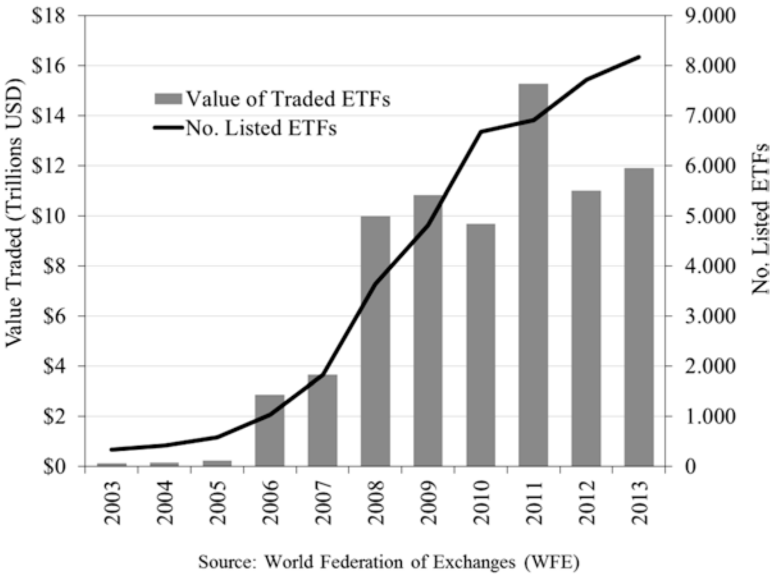


Fig. 14.5 Illustrated growth in the number of listed ETFs as well as the value of ETF trading on a global basis

alternative to a traditional index mutual fund that invests its assets in the constituents of a particular index in the appropriate proportions.

An index ETF is created when an institutional investor deposits securities into the fund in return for a fixed amount of shares or creation units. Investors may buy or sell fund shares on the stock exchange in a process identical to the purchase or sale of any other listed stock.

Net Asset Values—An ETF's net asset value (NAV) represents the value of all fund assets less the value of liabilities, divided by the number of shares outstanding. Investors may invest or redeem shares in an Open-End mutual fund at the NAV. ETFs are transacted at market prices driven by supply and demand considerations. Thus, the price of an ETF will typically resemble, but is nonetheless independent of, the underlying NAV of the fund.

Unlike closed-end ETFs, the shares for an index ETF may be created or redeemed on a daily basis by authorized participants. Institutional investors may redeem large share blocks "in-kind" if there is a gap between NAV and the market price of the fund. These arbitrage opportunities typically create sufficient demand to minimize the gap between NAV and market values.

ETF Legal Structure—There are three main legal or regulatory structures that have shaped the development of index ETFs described as follows.

1. Exchange-traded open-end index mutual funds are registered under the SEC Investment Company Act of 1940 (the '40 Act). Dividends are reinvested in the fund on the day of receipt and paid out quarterly in cash. Funds generally are allowed to use other derivatives and may generate income from security lending activity. Although there is no minimum amount an investor must purchase or sell, institutional investors may create or redeem shares in-kind in large blocks of perhaps 50,000 shares. Prominent examples of this structure include the Select Sector SPDRs and iShares family of ETFs.
2. Exchange-traded unit investment trusts are registered under the '40 Act and must fully replicate their benchmark index. However, diversification rules in the 40 Act sometimes compel these funds to deviate from an exact replication of index constituency and proportions. The '40 Act requires that no fund may invest more than 25 % of its assets in any single issuer. For non-diversified funds, the aggregate limit is 50 % of the total fund assets. To the extent that some indexes carry constituents at weights in excess of this threshold, several funds statistically optimize their holdings to reflect the index weightings. Dividends are not reinvested but rather are paid out quarterly in cash. Although there is no minimum amount an investor must transact, institutional investors can create or redeem shares in-kind in blocks of a specified number of shares (e.g., 50,000 shares). Examples include QQQs ("Qs"), DIAMONDS, S&P 500 SPDR, and S&P 400 SPDR.
3. Exchange-traded grantor trust—This type of fund is not registered under the SEC Investment Company Act of 1940, although it is similar to actual ownership of the underlying shares of the fund. The fund composition does not change, except to reflect corporate actions. These funds may be redeemed for the underlying securities, and investors have voting rights to the underlying securities.

Dividends are distributed directly to the shareholders and are not reinvested. Fund shares typically may be purchased/created and sold/redeemed in 100-share lots. Examples include Merrill Lynch's HOLDRs product line.

ETF Sponsors—Many entities are involved in the organization and operation of an ETF. But the primary role falls to the ETF sponsor. There are too many ETF sponsors to list but major players include Black Rock BGI and their iShares product line; Street Global Advisors (SSGA) and its SPDRs and streetTRACKS ETFs; Vanguard; Invesco which sponsors the PowerShares line; ProFunds which operates ProShares ETFs.

The ETF sponsor must license indexes from index publishers (as necessary) for use in the context of an ETF, establish the legal infrastructure of these products, engage custodians to carry the creation units, arrange for trading to be conducted on exchanges.

Types of ETFs—There are several types of equity-based ETFs as described below (Table 14.2).

1. Index ETFs—Most ETFs are constructed as index funds that hold securities and attempt to replicate the performance of a stock market index. Some ETFs invest 100 % of their assets proportionately in the securities underlying an index. Other index ETFs use representative sampling, investing 80 to 95 % of their assets in the securities of an underlying index, investing the remaining 5 to 20 % of their assets in other holdings, such as futures, option and swap contracts, and securities not in the underlying index, that the fund's adviser believes will help the ETF achieve its objectives. For index ETFs that invest in indexes with thousands of underlying securities, some index ETFs employ "aggressive sampling" and invest in only a tiny percentage of the underlying securities.
2. Actively Managed ETFs—Actively managed ETFs have been offered since 2008. They are transparent, disclosing current security holdings on websites daily. The fully transparent nature of existing ETFs implies that an actively managed ETF is at risk from arbitrage activities by market participants who might "front-run" their trades. Actively traded equity ETFs have addressed this problem by trading only weekly or monthly. But actively managed ETFs have received a lukewarm response to date. Reasons include concerns regarding front-running, the time needed to build performance records, and the failure of actively managed ETFs to give investors new ways to make hard-to-place investments.
3. Static ETFs—An exchange-traded grantor trust share represents a direct interest in a static basket of stocks selected from a particular industry. The leading example is found in the Merrill Lynch Holding Company Depository Receipts (HOLDRS) product line. HOLDRS are neither index funds nor actively managed. Rather, the investor has a direct interest in specific underlying stocks. While HOLDRS have some qualities in common with ETFs, including low costs, low turnover, and tax efficiency, many observers consider HOLDRS to be a product apart from ETFs.
4. Bull & Bear Leveraged ETFs—A leveraged ETF represents an actively managed ETF that attempts to achieve returns that are more sensitive to market movements than a non-leveraged ETF. They are marketed as either bull or bear funds.

Table 14.2 Top 20 equity-based ETFs ranked by value of ADV (Source: Bloomberg)

	Exchange Traded Fund (ETF)	Ticker	Category	Expenses (%)	AUM (000) 5/16/13	ADV in 3 Mths Ending 5/16/13 (000)
1	SPDR S&P 500	SPY	Large Cap Blend Equities	0.09%	$156,115,405	$20,914,460
2	iShares Russell 2000 ETF	IWM	Small Cap Blend Equities	0.24%	$25,931,472	$5,515,946
3	QQQ	QQQ	Large Cap Growth Equities	0.20%	$39,966,755	$3,863,412
4	iShares MSCI Emerging Markets ETF	EEM	Emerging Markets Equities	0.67%	$37,596,854	$2,897,685
5	iShares MSCIEAFE ETF	EFA	Foreign Large Cap Equities	0.34%	$55,157,005	$1,261,354
6	Dow Jones Industrial Average ETF	DIA	Large Cap Value Equities	0.16%	$11,029,882	$1,012,321
7	Energy Select Sector SPDR	XLE	Energy Equities	0.18%	$11,401,718	$985,842
8	Financial Select Sector SPDR	XLF	Financials Equities	0.16%	$17,563,778	$935,301
9	iShares MSCI Brazil Capped ETF	EWZ	Latin America Equities	0.60%	$4,451,373	$920,706
10	iShares China Large-Cap ETF	FXI	China Equities	0.73%	$4,796,730	$885,246
11	Market Vectors TR Gold Miners	GDX	Comm Producers Equities	0.52%	$7,556,463	$754,681
12	Core S&P 500 ETF	IVV	Large Cap Blend Equities	0.07%	$56,810,886	$713,879
13	Emerging Markets ETF	VWO	Emerging Markets Equities	0.15%	$44,562,230	$701,904
14	Nasdaq Biotechnology	IBB	Health & Biotech Equities	0.48%	$4,644,640	$591,918
15	Health Care Select Sector SPDR	XLV	Health & Biotech Equities	0.18%	$9,321,823	$590,047
16	Daily Small Cap Bull 3X Shares	TNA	Leveraged Equities	0.95%	$1,009,667	$572,246

(continued)

Table 14.2 (continued)

	Exchange Traded Fund (ETF)	Ticker	Category	Expenses (%)	AUM (000) 5/16/13	ADV in 3 Mths Ending 5/16/13 (000)
17	Ultra S&P500	SSO	Leveraged Equities	0.95%	$3,219,300	$566,593
18	Industrial Select Sector SPDR	XLI	Industrials Equities	0.18%	$9,899,601	$558,981
19	Utilities Select Sector SPDR	XLU	Utilities Equities	0.18%	$6,368,558	$533,034
20	SPDR MidCap Trust Series I	MDY	Mid Cap Blend Equities	0.25%	$15,059,994	$419,273

A bull ETF fund might attempt to achieve daily returns that are 2× or 3× more pronounced than the subject index while a bear ETF may attempt to achieve returns that are negative 2× or negative 3× as sensitive as the subject index, i.e., it will profits in a bear market. SEC has issued guidance to ETF providers that it will not approve ETFs going beyond 3× leverage, either long or short.

To achieve the desired leverage level, ETF providers actively rebalance the portfolio each day by buying or selling assets to maintain the constant daily exposure for the next trading day. There are several ways of achieving this goal. The manager can trade futures based on the same underlying index, or on a highly correlated index, or trade swaps in the OTC market. Given the nature of the rebalancing exercise, there is an embedded negative volatility exposure in leveraged ETFs.

14.1.7 Over-the-Counter Equity Swaps

Equity-based derivatives are traded on organized futures exchanges in the form of stock index futures and options on futures and on organized stock option exchanges based on individual equities as well as stock indexes. But there is a third venue for equity-based derivatives—the over-the-counter (OTC) markets. OTC equity instruments generally come in two varieties—swaps or forwards as well as options (Table 14.3).

An equity swap or forward contract essentially operates much like a stock index futures contract. It is typically cash-settled based on the value of an index on some future specified date. Unlike typical stock index futures, however, equity swaps are frequently settled vs. the total return (TR) version of a stock index, i.e., an index reflective of the price fluctuations plus the accrual of dividends associated with the index.

But the largest segment of the OTC equity derivative market is found in the form of equity options. These operate much like options traded on futures or security exchanges. Like equity swaps, however, they are frequently settled in cash by reference to the TR version of the index or instrument in question.

Table 14.3 Outstanding notional value of OTC equity swaps during the past several years

	Dec-2009	Dec-2010	Dec-2011	Dec-2012	Dec-2013
Forwards and swaps	$1652	$1828	$1738	$2045	$2277
Options	$4285	$3807	$4244	$4207	$4283
Total	**$5937**	**$5635**	**$5982**	**$6251**	**$6560**

Source: Bank for International Settlements (BIS)

OTC equity instruments are traded via a loose network of broker-dealers and their customers. While electronic matching platforms have become available in the context of many OTC derivatives, it is still most common to trade an OTC equity swap on a voice (or telephone) brokered basis. Because these contracts are still frequently negotiated and executed on a bilateral basis between a broker/dealer and customer, they may be highly customized to fit the needs of the customer, even if the broker/dealer subsequently lays off or hedges his risk using more standardized and liquid structures. The ability to create, account for, and manage the risks of customized products is facilitated, of course, by the availability of cheap computing power.

OTC instruments historically were unregulated in the United States. The Dodd-Frank financial reform legislation of 2010 generally brought swaps under the regulatory umbrella of the CFTC or SEC. But, as of 2014, there has been no regulatory mandate to require clearance of equity swaps or OTC equity options through a central counterparty clearing facility, as has been the case in interest rate and credit default swaps.

14.1.8 Securitized Derivatives

Securitized derivatives represent unique or "one-off" trading vehicles that are generally highly customized to serve specific trading or risk-management purposes. In that respect, they may be compared to customized OTC swaps. But securitized derivatives are typically issued by a broker/dealer or financial institution and then traded through an exchange. As a result of the customized nature of these products, they are typically of limited issue size. These products are quite popular in Europe and elsewhere in the world but have not significantly penetrated the US markets. Although the possibilities for design innovations are endless, they generally constructed as either certificates or covered warrants.

A certificate mimics a specified index or asset. They generally apply some degree of leverage although so-called "investment certificates" may be designed as unlevered vehicles. Covered warrants represent a form of call or put option. While it is common to create a simple covered warrant, complex option structures or strategies may be embedded in the product structure. Warrants may be based upon individual assets, portfolios, or indexes. While the number of listed securitized derivatives continues to grow, values traded have dwindled in recent years as illustrated in Fig. 14.6.

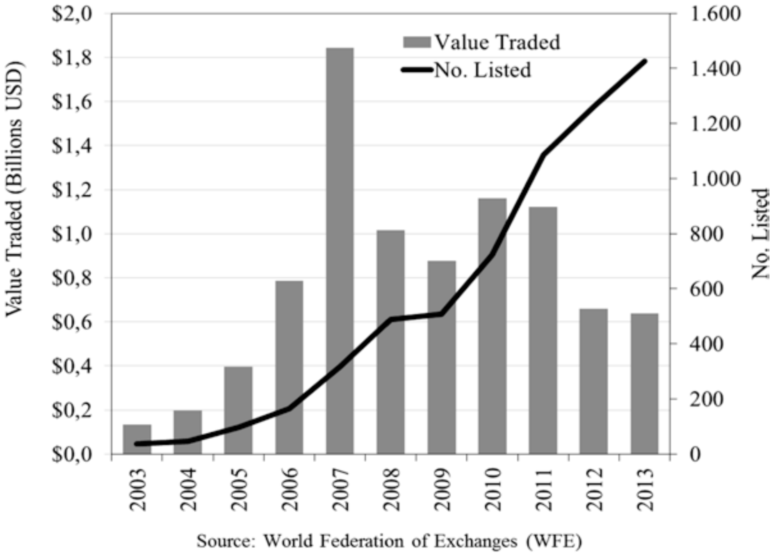


Fig. 14.6 Securitized derivatives activity

14.1.9 Emerging Innovations

Innovation is the constant pursuit of the financial engineers who busily create new products to address market demands. Two of the most interesting new concepts to emerge in recent years include volatility and dividend-based products.

Volatility Products—Volatility refers to the movements of a stock or stock index in an absolute sense. Volatility represents price moments in a directionally neutral way, without referring to bull or bear trends. A market may be volatile when it is rallying or breaking, provided it is moving a lot. Volatility may be measured as the annualized standard deviation or variance of daily price returns in an equity or equity index. Alternatively, volatility may be measured by reference to the implied volatility (IV) of an option[2].

Since the late 1990s, a specialized market in cash-settled equity volatility swaps has evolved. These swaps typically reference the variance associated with an

[2] One may utilize mathematical option pricing models to calculate the “fair value” of an option premium (Prem) as a function of the market price (M), exercise or strike price (E), term until expiration (t), volatility (v) short-term interest rates (r) and dividends (d).

$$\text{Prem} = f(M, E, t, v, r, d)$$

But volatility, as a rather abstract concept, is not readily observable like the other variables. However, one may easily reference the prevailing option premium in a liquid market. Thus, one may solve an option pricing model to find the “implied volatility” (IV) or volatility implicit in the prevailing option premium as a function of the other known variables.

$$\text{IV} = f(\text{Prem}, M, E, t, r, d).$$

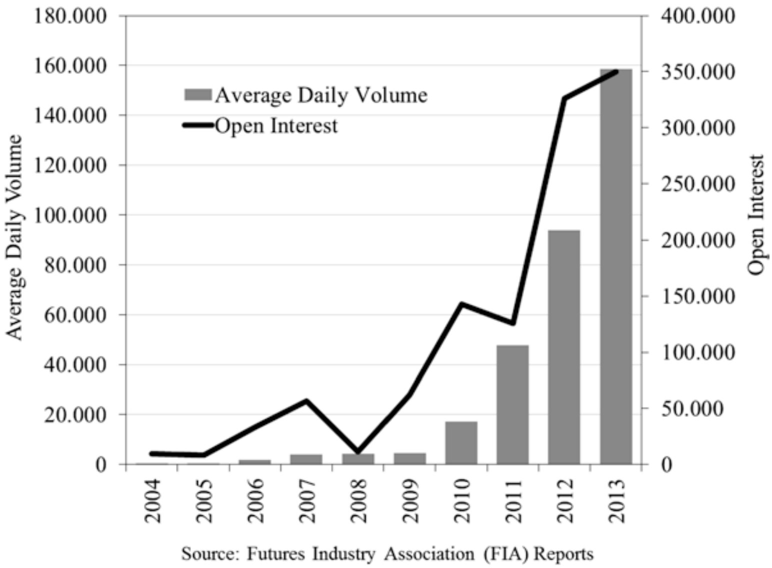


Fig. 14.7 CBOE VIX futures activity

equity or equity index observed over a specified time interval. Variance has certain mathematical properties which facilitate linkage with option markets. But variance is explosive in the sense that it can advance exponentially, as it represents the square of standard deviation, exacerbating risk and margin requirements.

The S&P 500 Volatility Index (VIX) measures volatility as an index or average IVs of liquid stock options traded on CBOE. Very actively traded futures and option markets are based upon the VIX as illustrated in Fig. 14.7. This has also prompted development of successful ETFs that hold VIX futures.

Dividend Products—Dividends have likewise become the focus of equity-based derivatives in recent years. Actively traded OTC swaps, futures and options are based upon the realized dividend streams associated with popular stock indexes. While these markets thrive in Europe, regulatory issues impeded the development of dividend-based futures in the USA for many years. Futures based on dividend indexes were recently launched in the USA and are gaining traction.

14.1.10 What's Next?

Equity derivatives will continue to evolve and improve in pursuit of more effective ways to serve stock traders. While we can only speculate on what direction these new developments may assume, we may observe that competition amongst exchanges, ETF sponsors, broker/dealers, and others to win marketshare and revenue will inevitably motivate the creation of both disruptive or radical, along with incremental or evolutionary, new products moving forward. This competition is facilitated by the availability of cheap but powerful computer technologies that support product proliferation with ever-increasing degrees of customization and sophistication. These products must, of course, conform to approve regulatory structures which likewise will continue to shape market development.

On the other hand, competition and customization may imply product proliferation and a possible fragmentation of the inherently limited liquidity pool. This may be magnified to the extent that regulatory "silos" sometimes encourage the creation of parallel derivatives markets serving roughly the same purposes but offered per different regulatory structures. Progress forward, therefore, requires a balance between this customization to serve very specific needs with the need to provide liquidity, within the framework of current regulatory structures. In the final analysis, of course, the ability to attract participation and achieve a critical mass of liquidity remains the ultimate test for any marketplace.

Chapter 15
From the End of Bretton Woods to the Global Financial Crisis: 40 Years of Turbulence

Hugo Bänziger

15.1 Introduction

In modern history books, the expression financial crisis is a well-known term. At the time of the Great Depression a staggering 40 % of countries saw their banking system fail. Nearly 80 years later, at the peak of the Great Financial Crisis, the number reached 30 % again. However, from 1945 to 1973, under the Bretton Woods agreement, the global financial system was quite stable. But with the end of Bretton Woods, financial turbulence returned. Even Switzerland, Canada, and Australia, countries recognized as stalwarts of financial stability, experienced a banking crisis.

The Great Financial Crisis in 2008 did major damage to the world's financial system. The IMF estimated total losses to reach USD 2.0 trillion. Taxpayers had to spend more than USD 500 billion to rescue banks. Far worse, the Great Financial Crisis pushed the world into a deep recession, with the fallout still discernible in many countries. The world's GDP dropped by USD 3.3 trillions (−5.3 %) and world unemployment increased by over 20 million. This spillover effect highlights the importance of maintaining financial stability as a public policy goal. It is thus important to ask whether the end of Bretton Woods was linked to the increase in financial instability, or whether it was perhaps the cause. This chapter is written from an economic policy perspective and presents a brief description of the Bretton Woods system followed by a short description of the many financial crisis since 1973. The text also reflects on the impact of globalization, information technology, rise of institutional investors, and deregulation across the financial sector. Finally, it addresses the ever-growing complexity of the financial system and analyses whether the necessary reform process for banks and the financial infrastructure is complete.

H. Bänziger (✉)
Lombard, Odier & Cie, 11 Rue de la Corraterie, Geneva 1204, Switzerland
e-mail: h.banziger@lombardodier.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_15

15.2 Banks Under Bretton Woods

The Bretton Woods system was the result of sweeping political and economic reforms in the USA and parts of Europe from the Great Depression. The enormous structural changes imposed on banks were arguably more profound than the regulatory reforms we experience today.

On 6 March 1933, as US banks were on the brink of failure, the newly inaugurated President, Theodore Roosevelt, suspended all banking transactions and declared a week-long bank holiday. Banks could only reopen and regain their access to the Federal Reserve window if they were judged solvable.

Three days later, he introduced the Emergency Banking Act. The bill was submitted under such urgency that only a single copy was available. Henry Steagall, the Chairman of the House Banking Committee, was obligated to read it aloud on the floor of the House. The Act gave Roosevelt the power to regulate the Federal Reserve System and to recapitalise or resolve banks.

The legislative response to the banking crisis continued at a fast pace. On 27 May 1933, the US Congress approved the Securities Act, regulating primary securities markets. The bill required issuers of stocks and bonds to register and disclose standardised financial information. The Act was the first step towards the financial transparency from which we benefit today.

A month later, the Banking or Glass-Steagall Act was established. Its main goal was to split investment banking from commercial banking to prevent losses spilling over from securities trading into deposit taking. The Act's regulation Q governed banks' liabilities, eliminated interest payments on checking accounts, and capped rates on deposits. The Act also created the Federal Deposit Insurance Corporation (FDIC) to insure all deposits up to USD 5000. It established large exposure limits and laid the basis for regulatory reporting and supervision by the Federal Reserve.

The last big reform packages were implemented in 1934 and 1935. The Securities and Exchange Act of 1934 created the Securities and Exchange Commission (SEC) and regulated secondary trading of stocks and bonds. It also governed the operation of exchanges. The Banking Act of 1935 completed the regulatory overhaul. It restructured the Board of the Federal Reserve System.

The American banking system thus changed dramatically between 1932 and 1935. Lending and deposit taking became the main business of banks with margins tightly regulated. Working capital loans and mortgages now dominated the balance sheets and leverage was substantially reduced. Capital market operations underwent dramatic change as well. Investment banking units had to split from their parent banks and became independent broker/dealers, mostly organized as private partnerships. The issuance of securities now required a standardised prospectus and trading on exchanges became mandatory.

Whilst the USA was the most radical reformer, European countries established their own banking laws in the Great Depression. They mostly stipulated minimum capital requirements, liquidity ratios, large exposure rules, cross-border restrictions, and rules for governance and audit. However, Europe left the universal banking model intact. Some nations tasked the central bank with banking supervision. Others created new government agencies.

A few years after these fundamental reforms, with the outbreak of the Second World War, the financial system in every part of the world assumed a new role. Banks became instrumental in financing the war effort. War bonds were the primary instruments in Anglo-Saxon countries. In the rest of the world, banks invested customer deposits in government loans and certificates. Assets related to the war effort now dominated the balance sheets.

In 1944, with victory in sight, the Allies started planning the post-war financial system. Representatives from 44 countries gathered in the small town of Bretton Woods, New Hampshire, to shape the new financial order. There were to be three pillars. The first was the promotion of free international trade, a concept mostly abandoned during the depression years. It resulted in the General Agreement on Tariffs and Trade (GATT) in 1947. The second pillar established fixed exchange rates with capital controls and created the International Monetary Fund (IMF) to provide funds if members experienced balance of payment problems. The third pursued the reconstruction of the world economy from the ruins of war, establishing the International Bank for Reconstruction and Development (IBRD), which is now part of the World Bank.

Under the Bretton Woods system, all currencies were pegged to the US dollar at $35 per ounce of gold. It thus maintained the semblance of the old gold standard, an intuitive solution when the USA held three-quarters of the global stock of gold. The dollar emerged as "the" international currency. Theoretically, the fixed pegs could be adjusted in order to correct any "fundamental disequilibrium" in the balance of payments, but only France (devaluations in 1949, 1958, and 1969), Germany (appreciations in 1961 and 1969), and the UK (devaluations in 1949 and 1967) ever adjusted the exchange rate. Combined with capital controls, this system eliminated exchange rate risk.

The end of the Second World War was for the USA and the rest of the world a very different experience. The USA exited the war with an intact and very modern industrial base. This allowed the country to switch quickly to consumer goods production and become a main exporter. The resulting tax receipts enabled the US Government to reduce its war debt within 10 years, from a record peak of USD 120 billion to less than half. Other countries were not as fortunate. With their industrial basis either destroyed or severely outdated, they faced the harsh reality of large balance of payment deficits, inflation, and currency devaluation. The German and Japanese banks, which had little more than war-related assets on their balance sheet, were essentially bankrupt. It took billions of dollars of bilateral US loans, the Marshall Plan, and eventually the 1948 currency reforms in Germany and Japan, to allow the global economy to grow again. The low peg rate of both DM and Yen favoured the building of export-led economies in both Germany and Japan.

Banks in Europe and Japan were in consequence very different from their predecessors prior to the Second World War. With the war, capital markets had virtually disappeared. They were now the sole source for working capital and long-term investment financing. Their funding relied on domestic deposits and benefitted from high domestic savings ratios. International bank lending remained at very low levels. As the US peers, European and Japanese banks were tightly regulated. Interest rates were set by the central bank and governments often directed their lending activities. There was little competition between banks, which began to resemble large government agencies. Not surprisingly, these banks were neither efficient nor innovative.

Being tightly controlled or nationalised, banks lost their skills in making risk-reward decisions. With Bretton Woods limiting the cross-border capital flow and keeping exchange rates fixed, banks also lost their skills in international lending and managing market risk. Since capital markets were severely damaged by the war, equity and corporate bond underwriting remained low in the post-war years. Investment banking became an insignificant activity. As a result, banks were not very complex and easy to understand. They carried mostly domestic credit risk, almost no market, and very little liquidity risk. Not surprisingly, this long period of stability gave birth to several generations of bankers whose understanding of financial risk was limited.

15.3 Bretton Woods Ends …

By 1960, Japan and Germany were back again on world stage as major exporting countries. International trade was gaining pace and restrictions on international capital flows were being relaxed. At the same time, inflation reappeared in America. President Johnson's spending for the "Great Society", together with the spiralling cost of the Vietnam War, led to continued budget and balance of payment deficits. Rapidly growing money supply and diminishing gold reserves put pressure on the gold-pegged US Dollar. Despite numerous government interventions through the London Gold Pool, the USA abandoned in 1968 the US Dollar's 25 % minimum gold coverage. France and other countries, predictably, began to increase their physical holding of gold. Within months the London Gold Pool collapsed. By 1971, US President Nixon unilaterally terminated the Dollar's gold convertibility. Two years later, Japan and Germany let their currencies float. Since most banks were only active in their home market, the end of the Bretton Woods system seemed to have no discernible impact on financial stability. But this was going to change.

15.4 … And 1973 Changes Everything

Ever so often, there is a year in history, which stands for significant change. 1973 was such a year. The Black-Scholes model for options pricing was published and became widely used by market participants creating a common basis for volatility trading. That very same year, the Chicago Board Options Exchange began trading.

1973 also saw the world's first mobile phone from Motorola and optical glass fibres were invented that year. Both technologies would become the backbone of the revolution in communication that has continued ever since. They would not only connect the world. They would connect the world's markets.

Innovation in information technology continued a pace. Whilst IBM had offered computer-based services for processing equity trading as early as in 1959, IT remained confined to the back offices of banks. By 1973 however, this was changing. Desktop machines such as the IBM S/370 and HP 9800 became available and quickly spread through banks' front and middle offices. For the first time, trading records could be maintained electronically. Less noticed was the installation of Citibank's first ATM, which would revolutionise retail banking and create new business opportunities around the globe.

1973 was also a year of geopolitical shocks. The Watergate scandal was in full swing, eventually leading to the resignation of US President Richard Nixon the following year. More importantly, Israel's victory in the Yom Kippur War in October 1973 triggered an oil embargo by the Organization of Oil Producing Countries (OPEC) against the USA and all countries that had supported Israel. Within days, oil repriced sharply and reached a peak in today's Dollar, which would not be reached again before 2008. The world of cheap energy had gone. Europe experienced its "car-free" weekends. The world dropped into its first global recession since the Second World War.

15.5 Animal Spirits Return

In this turbulent year, Germany and Japan's move to floating exchange rates made headlines but it was not really clear what the end of Bretton Woods would mean. But warning signs popped up early. Herstatt Bank, a German medium-size lender, collapsed in 1974 due to its unhedged currency exposure. It had not adjusted to floating exchange rates. The name "Herstatt" became a synonym for foreign exchange risk.

Over the next two decades, liberated from the constraints of Bretton Woods, five large forces would reshape the world of finance. The first was the return of international banking, which had been subdued after 1945. The sharp increase of oil prices flooded the Middle East with large amounts of US Dollars. For political reasons, oil producers preferred to keep these petrodollars outside the USA and placed them primarily with banks in the City of London. It was the beginning of the Eurodollar market, which would revitalise London as a global financial centre. Most petrodollars were recycled via syndicated loans to sovereign borrowers in emerging markets. Latin America alone quadrupled its international borrowing in the years following the end of Bretton Woods. A smaller part of the petrodollars would flow into the new Eurodollar bond market. Its development was so dynamic that within 5 years the first dedicated Eurobond House, Credit Suisse First Boston, was established. The City of London became the global hub for foreign exchange trading and the Eurodollar market. Attracted by these business opportunities, banks from around the world established their presence in London. New branches and subsidiaries were opened. With little experience in international lending, the Bretton Woods generation of bankers became the underwriters of syndicated loans and bonds. “A country does not go bankrupt”, said Walter Wriston, Chairman of Citibank, famously. The sovereign debt crisis a few years later would prove otherwise. Cross-border lending became a permanent feature in international finance, but also a permanent source of troubles.

The second big market force was the rapid innovation in information and communication technology. Just 4 years after 1973, General Telephone and Electronics sent the first live voice traffic through a fibre-optic cable in California. By 1988, the first transatlantic fibre-optic cable was laid. Simultaneously, computers became smaller and easier to use. IBM introduced its very successful Model 5150, a personal computer, in 1981, equipped with the elegant spreadsheet application Lotus 1-2-3. Within a few years, Lotus 1-2-3 became the most widely used software in the financial industry, changing banking forever. Whilst it was perfectly possible to price cash flows of futures or swaps manually, the process was time consuming. By the time the result was calculated, the underlying prices had moved, thus making the result irrelevant. Lotus 1-2-3 solved that problem. With its Macro programming function, cash flow calculations could be done in real time. In one swift action, financial products called derivatives, which derived their value from underlying instruments, could be valued and traded like the underlying product. The rapid increase in IT processing capacity also resulted in a steep decline of transaction costs. For the first time in history, investors could trade frequently without sacrificing yield. The active management of investment portfolios became possible. As a result, trading volumes increased rapidly.

The combination of low transaction cost and real-time pricing resulted in another important development: the emergence of a global derivatives market. It was not by accident that SOFFEX, the world’s first electronic options and futures trading exchange, commenced operations in 1988. Today, under the name EUREX, it is one of the largest derivative exchanges in the world. At the same time, these advances also laid the ground for over-the-counter (OTC) market. J.P. Morgan, Bankers Trust, and Credit Suisse played a dominant role in its development in the early 1990s. Within a few years, derivatives linked all major capital markets. By 1996, McKinsey wrote in the study “Markets Unbound” that eventually everything would become tradable. As with every innovation, original margins on derivatives were high. By 1990, banks charged 8 bp for an interest rate and 20 bp for a currency swap. Derivative trading desks were established in every bank. The volume of derivative trading took off on a scale that frightened many. But more supply also meant more competition. By the time of the Global Financial Crisis, margins had dropped by more than ten times.

Last but not least, the global fibre-optics network now linked markets and exchanges around the globe. Price differentials between markets could be arbitraged. Today, it is a faint memory that IBM issued bonds in Japan cheaper than in the USA because Japanese investors considered IBM to be a blue chip name for much longer. Arbitrage eliminated these price differentials. The world became a truly global market. It also made banks much more complex. Far-flung operations around the globe and high trading volumes required significant IT investments. It was not by chance that the financial sector replaced defence and aerospace as the single biggest IT customer.

The third large force at work was the emergence of institutional investors. By 2002 they became the largest holders of financial assets. Whilst the insurance industry had its roots in the nineteenth century, mutual funds and pension plans were much younger. Before 1960, they were hardly noticed. Mutual funds were a market response to the low returns bank customers could earn on their regulated deposits, and pension plans to the ageing of society, which made government pay-as-you-go pension systems unsustainable. Both would be inconceivable without information technology and the efficiency gains it yielded. The mutual funds industry started slowly. By 1970, it had only USD 48 billion under management. By the end of 2011, the number had increased to USD 24 trillion. Pension plans saw an even faster development and by the same year managed USD 31.5 trillion of assets. Together, insurance companies, mutual funds, and pension plans managed around USD 80 trillion of assets, far exceeding the USD 26 trillion bank deposits globally. The rise of institutional investors had far-reaching consequences. With their statutory return requirements, they established capital market discipline. Institutional investors have to earn a return on investment. The positive economic impact of institutional investors is not to be underestimated. Asking for "shareholder value", they reallocate capital from underperforming businesses to more profitable investments. As such, they were at the core of the industrial revitalisation of the 1980s and 1990s.

For banks, this transition proved to be difficult. Being run like government agencies or as instruments of public policy, banks were not the most profitable organisations. To maintain their access to capital markets, they had to improve their profitability. Banks thus embraced the new business opportunities, which the globalisation of finance offered, wholeheartedly. But these opportunities were a double-edged sword. The reward came associated with risks. Any institution can improve its return by increasing its risk profile, specifically, when the new business activities do not require any new capital, as was the case under early regulatory rules. Improving returns by increasing efficiency is much harder. The banks' preference for increasing their risk profile was thus entirely rational.

But being subject to capital market discipline had consequences, which few noticed at the time. As long as banks perform, they attract investors easily. However, in times of stress, when risks and losses crystallise, the story is different. Investors resent throwing good money after bad money. They would inject equity or rollover debt when presented a credible new business plan but otherwise stay on the sideline. Recapitalising or refinancing a financial institution under duress would prove to be challenging. Almost through the backdoor, Schumpeter's principle of "creative destruction" found its way back into the financial industry. Finance was governed again by two contradicting public policy principles: financial stability and capital market discipline. In the absence of bank-specific resolution tools to unwind failed banks, it was difficult to see how these two policy principles could reconcile.

The forth force shaping finance after Bretton Woods was the international expansion and consolidation of banks. As currency and capital controls were abolished, banks pursued new business opportunities abroad, which were often more profitable than operations in their regulated home markets. US banks were the first to take advantage. They substantially expand their branch network abroad with London as the key hub, followed by Frankfurt, Tokyo, Singapore, and Hong Kong. Most of the world's banks followed. Citibank is a good example. It has now more than 9000 branches in 44 countries. Santander, a Spanish lender, earns just 13 % of its profit in its home market. After the global expansion, a wave of domestic mergers followed with the aim to achieve better economies of scale. When Hong Kong and Shanghai Bank (HSBC) acquired Midland Bank in 1992, the period of cross-border acquisitions began. They brought together institutions with very different cultures. Today's large banks in the USA, Europe, and Asia are mostly the result of mergers between rival institutions. Not only do they run large operations in many different jurisdictions, but they also diversified. A large bank typically comprises today commercial and investment banking operations, retail and transaction banking, and private wealth and asset management. Several banks also have large insurance companies within their fold. The growing complexity puts enormous pressure on their governance. The integration of complex IT systems proved challenging and time consuming. It is fair to say that mergers slowed down the development of modern management information systems—a field where banking considerably lag the manufacturing industry. By the end of the 1990s, banks were large and complex but governed and supervised by institutions from the time of Bretton Woods.

Liberalisation was the last of the five big forces to shape finance. Less than a decade after abolishing the Bretton Woods system, deregulation of domestic markets followed. The regulatory framework of the Great Depression was stifling growth at a time when growth rates started falling. The USA removed Regulation Q, which capped interest rates on bank accounts, from 1981 to 1986. Business restrictions for America's savings and loan associations, the so-called Thrifts, were lifted during the same years. In 1984, the barriers to interstate banking were eliminated. The Glass-Steagall Act, which separated investment from commercial banking, fell de facto by 1989, when the Federal Reserve allowed Credit Suisse, a Swiss universal bank, to rescue the insolvent US investment bank First Boston. It would take another 10 years to formally lift the Act. But after 1989, the Glass-Steagall Act was no longer enforced. A similar development took place in the City of London. In 1986, the UK Government lifted with the "Big Bang" the traditional separation between banking activities, allowing commercial banks to expand into broking and dealing. During these years, Europe followed its own path with the Maastricht agenda. In 1992, the Maastricht Treaty introduced the four liberties for goods, capital, labour, and services. In essence, a banking license in one country entitled a bank to offer the full range of products in any EU country. Deregulation was not unilaterally welcome. Many considered the policy as a dogmatic result of the Reagan-Thatcher years. But it was a pragmatic macroeconomic policy at a time of lacklustre economic growth. In the EU Commission's view, a working financial market contributed 1 % to GDP growth per year.

Liberalisation was not limited to regulatory rules alone. In most countries, trading of securities on exchanges was mandatory. These rules were established to enforce trading discipline and prevent fraud. But the progress in information technology enabled banks to pool client orders and net them internally to minimise exchange-trading fees. In essence, they formed invisible internal pools of liquidity, so-called dark pools. With ever-increasing trading volumes, dark pools became a profitable business. Understanding their customers' trading activities also allowed banks to build proprietary trading desks. Not surprisingly, securities trading on exchanges dropped to a fraction of the total volume.

A second, almost unnoticed, development took place in New York in 1971, when the New York Stock Exchange dropped the requirement for its members to be organised as partnerships. As a result, over the next two decades, all investment banks converted to joint stock companies. The prudent and conservative partnership structure was replaced by a new, more aggressive corporate culture.

15.6 Corporate Governance Outpaced

Despite the fast pace of change throughout the 1980s and 1990s, corporate governance had changed little. Representatives of industrial companies or government officials continued to dominate the boards. Promotions to executive positions were mostly based on seniority. Management information was often not more than the statutory accounts of legal entities. As in the calmer times of Bretton Woods, decision-making was highly centralised and most issues escalated to the top. But banks had grown in size and complexity. The management process in many institutions became seriously stressed. Modern management information systems had yet to be introduced. Most executive meetings were day-long affairs and executives had to digest yards of briefings and submissions for each meeting. In most institutions, a daily profit and loss statement did not exist and issues were decided case by case without any consideration on how they might affect the portfolio of risks. Bank executives were very busy but often found little time to understand the dynamics and risks of their new businesses. Not surprisingly, they were often surprised when things went wrong.

The Latin American Debt Crisis and the stock market crash of 1987 were indeed sobering events. They revealed how unprepared banks were in managing risk. Whilst lending officers started to use credit ratings to assess the risk of their loan books, economists and mathematicians on the trading floors began applying statistics to better understand market risk. Eventually, value at risk (VaR) was developed. Data sets with historic prices or Monte Carlo simulations were used to extrapolate extreme price movements. VaR enabled banks to predict "worst-case" losses and monitor them against new limits. In 1990, J.P. Morgan became the first bank to use VaR across its trading desks. The CEO insisted on a daily profit and loss statement including market risk numbers within 15 min of the closing of the New York Stock Exchange.

At the same time, derivative exchanges faced problems with calculating collateral margins. Monthly accrual accounting was of no help. A real-time profit and loss statement was necessary to make accurate margin calls. The concept of marking a book to market by using the current market price for each position was born. It crossed quickly over to leading OTC derivatives houses, which used it to value their trading books. By 1993, the American accounting standard setters approved the "mark-to-market" or "fair value" concept for equity and debt securities trading by adopting rule FAS 115. The new rule also allowed the netting of positions within the same trading book. Hedging of risk had found its official approval. "Fair value" accounting was quickly taken up around the globe. A key moment was the violent turn of the USD yield curve in spring 1994, which caused significant losses for banks, which had retained accrual accounting for trading. The concept also found its way into the International Financial Reporting Standards (IFRS), the European equivalent to FAS. With the adoption of "fair value" standards, an important public policy decision was made. Banks were permitted to put any transactions into the trading book, if they "intended" to trade. But there have been few truly liquid instruments, even in today's markets. The biggest foreign exchange pairs are liquid, the largest government bond markets, maybe the top 200 global stocks and some commodities such as oil or gold. All other financial instruments are semi-liquid or illiquid. Any bigger transaction moves the bid-offer spreads quoted by traders and market makers. "Fair value" thus opened the door for putting a whole range of less or ill-liquid products on the trading book with consequences, which became apparent many years later.

Central bankers observed the foray of banks into new business activities with a given dose of scepticism. Following the collapse of the Herstatt Bank, they established in 1975 the "Basel Committee for Banking Supervision" to monitor the health of the global banking system. But developing international standards proved to be difficult. Only when the Latin American debt crisis revealed the weakness of major international banks, things started to move. By 1982, the US Congress raised capital requirements for US banks and tightened their supervision. A year later, Paul Volcker, Chairman of the Federal Reserve, asked for an international capital accord. But it would take 5 years to establish it. By 1988, the world had finally a global capital standard, later called "Basel I". For the first time in history, a bank's minimum equity had to be based on the riskiness of its assets. But the accord covered only credit risk. Minimum requirements for marker risk still had to be developed. Liquidity requirements were not even discussed.

It would take another 8 years to complement the Basel I accord with capital requirements for market risk. By 1996, the Basel Committee for Banking Supervision (BCBS) had standardised the banks' VaR models, making it the basis for market risk capital. It was designed for risks in liquid foreign exchange and government bond markets, not for semi-liquid or illiquid instruments. But in a kind of mission creep, the new market risk capital standard was quickly adopted for all trading books. This design flaw would have colossal consequences. Based on the assumption that positions in a trading book could be sold quickly, market risk capital requirements were considerably lighter than traditional capital charges. But under "fair value", the intention to trade mattered, not the liquidity of the position. Thus banks had an incentive to put assets into the trading book. The door was open for arbitrage between credit and market risk capital requirements. Not surprisingly, leverage in the financial system started to increase. By the end of 2006, Citibank had USD 109 billion of risk-weighted assets for a trading inventory of USD 538 billion, and Deutsche Bank EUR 14 billion for EUR 596 billion. In a traditional accrual book, risk-weighted assets are closer to 100 % of the underlying asset values.

15.7 The Return of Turbulence

The rapidly changing environment did not only offer new business opportunities, but it also brought back the financial turbulences meant to be tamed by the 1930s' reforms and Bretton Woods. The first sign of trouble came from the recycled petrodollars, which found their way into emerging markets. In the 1970s, Latin America quadrupled its external debt to USD 315 billion. Most of these loans were made available by American, European, or Japanese banks, and invested in prestigious but commercially not always viable infrastructure programmes. The second oil shock in 1979 and the sharp tightening of US monetary policy in 1981 made these loans unsustainable. In 1982, Mexico and Brazil defaulted. Eventually, the debt of 16 Latin American countries had to be restructured and lost 1/3 of its value. Brady bonds, issued in 1989 and named after the US Secretary of the Treasury, crystallised these losses. The Latin American debt crisis had far-reaching consequences for the continent. Years of austerity followed the debt restructuring. Today Latin Americans call these years "the lost decade". The crisis also revealed the weak capitalisation of America's major banks. To prevent them from becoming insolvent, US regulators allowed them to delay the impairment of their Latin American loans. First loan loss provisions were only established in 1987, 5 years after Mexico's default. The regulatory forbearance weakened market discipline and may well have contributed to the excessive risk taking in later years.

In 1984, in the middle of the Latin American debt crisis, the USA saw the first bank bailout after the Second World War. Continental Illinois, a Chicago-based money centre bank, became insolvent due to large non-performing loans from its business with the Texan oil and gas industry. When depositors withdrew USD 10 billion within a few weeks, the Federal Reserve feared widespread contagion. Given the weakness of the US banking system after the default of Mexico and Brazil, Continental was rescued with an injection of USD 5.5 billion new equity and USD 8.0 billion of emergency loans. Shareholders were wiped out but bondholders—even on the holding level—were spared. It was to be the largest bailout until 2008. US Congress members coined the phrase "too big to fail".

Continental Illinois and the Latin American debt crisis were not the only problem the US economy faced. The loose monetary and fiscal policy of the Johnson era left the country with high levels of inflation. Only drastic actions of the Federal Reserve under its new chairman Paul Volcker brought it under control. In 1981, Volcker raised the federal fund rates quickly to a peak of 20 % and restored sound monetary discipline. The market upheaval was unprecedented. It not only triggered the Latin American debt crisis, but also undermined the business models of America's savings and loans institutions. Having a market share of 80 % of the USD 700 billion mortgage market, the Thrifts faced a unique problem. Whilst providing long-term, fixed-rate mortgages to their customers, they refinanced themselves with short-term deposits. When short-term interests went sky high, the Thrifts' net interest margin became negative. They paid more for their short-term deposits than they could earn on their mortgages. Rather than addressing the underlying asset—liability mismatch—US regulators allowed them to expand into general banking but without imposing the same supervisory standard, which was in place for other banks. Not surprisingly, many Thrifts expanded into very profitable, but also highly speculative real estate developments and shopping mall financing. Many of these investments were fraudulent. From 1980 to 1983, the Thrift's balance sheet increased by almost 60 %. The wake-up call came quickly. After 1986, one-third of the 3200 savings and loan institutions became insolvent. By 1989, the crisis had reached such an extent that US President Bush had to establish the Resolution Trust Corporation (RTC). The RTC's final bill amounted to USD 160 billion. US taxpayers had to cover USD 120 billion. The S&L crisis was as expensive as the Latin American debt crisis. Within one decade, the American financial system lost close to USD 250 billion.

Turbulence not only raised its head in America. The next country to be affected was Japan. The Japanese asset bubble emerged from a healthy, export-driven economy with high saving rates and almost unrestricted supply of money and credit. Liquidity was lavish in the late 1980s and supported by the Bank of Japan rather than curtailed. Japanese commercial banks, seeking to put their abundant deposits to work, encouraged their clients to borrow. The demand for both equity and real estate increased quickly with prices attaining their peak in 1988. The asset bubble imploded when the Bank of Japan raised interbank rates in 1989. The tumbling asset prices sharply increased the banks' non-performing loans and revealed the ultrathin capitalisation of Japanese banks. Several were at the brink of insolvency and could only survive with government support. Sanyo Securities, Yamaichi Securities, Long-Term Credit Bank, and Nippon Credit Bank had to be rescued. Many others were merged. This wave of after-crisis mergers shaped the large banking conglomerates, which dominate Japan today. The taxpayer bill summed up to Yen 9.3 trillion or USD 91 billion. More importantly, the collapse of the Japanese banking system led to two decades of anaemic growth and brought back deflation, something the world had last seen during the Great Depression.

Financial crisis began to erupt in other countries as well. Scandinavia saw its property bubble peak at the beginning of 1990. When it eventually burst, the Nordic governments were forced to guarantee the banks' deposits. Large loan loss provisions turned the banks' profit and loss statements deeply red. Nordbanken and Gotabanken, two Swedish banks, were nationalised to avoid a catastrophic collapse. Today, the Nordic bank resolution is considered as one of the most successful in history. However, Swedish GDP dropped by 5 % between 1990 and 1993 whilst unemployment increased sharply. The price of a financial crisis continued to stay high.

Four years later, the Tequila Crisis in Mexico reminded everyone how quickly capital could be pulled out of a country. Twelve years after its first default, Mexico, which had just signed the NAFTA free-trade agreement with the USA and Canada, needed the help of the US Government to avoid a financial meltdown. To recapitalise its broken banks, Mexico had to sell its three largest to Citibank, BBVA, and Santander. The Tequila Crisis came at the cost of losing ownership of the banking system.

In 1997, excessive private sector borrowing triggered the Asian debt crisis. With central banks running out of foreign currency reserves, the local currency pegs to the US Dollar had to be relinquished. 1998 also saw the default of Russia. External investors had lost their confidence in the Russian Government after the events in Asia and started selling the Russian currency. Within days, Russia's foreign debt could not be serviced any longer. The banking system in both Asia and Russia suffered heavily. Severe recessions followed.

Common to all those events between 1973 and 1998 was the fast transition from heavily regulated domestic banking to unregulated global markets and the rapid change in the global macroeconomic environment. Banks were quick to discover and take advantage of new business opportunities. But they were slow in identifying the related risks, capitalising them and in making the necessary adjustments to their corporate governance. The new risks remained mostly unmanaged. The new businesses undercapitalised. Even worse, "fair value" accounting and the new capital standards for market risk opened the door for regulatory arbitrage. The failure in the private sector had an equivalent in the public sector. The necessary evolvement of the regulatory framework came far too late. The Basel I Accord arrived only in 1988, and the minimal capital requirements for market risk only in 1996. Given that the new trading activities required no or little capital, the leverage in the financial system increased. All this happened at a time when the complexity of banking operations increased beyond what was commonly understandable. The analysis of annual reports of banks, which failed in one of the several crises, reveals that none had identified the risks they carried.

15.8 Catching Up with the Animal Spirit

The rapid succession of crises after crises caused considerable soul searching amongst central bankers. By 1999, they called for improving risk management in the financial industry. The same year, work began to overhaul the basic Basel I framework. But despite the well-known limits of the VaR model, market risk was excluded. Basel II was to focus on credit risk. As early as 2000, Myron Scholes, one of the fathers of the Black-Scholes model, voiced his concerns: "VaR, the product of portfolio theory, is used for short-run day-to-day profit and loss-risk exposures. Now is the time to encourage the BIS and other regulatory bodies to support studies on stress test and concentration methodologies". But his voice remained unheard. The fact that the bursting of the dot.com bubble did not cause serious losses in the US banking industry led many to believe that banks could well cope with market risk under the rules established in 1996. The bursting of the bubble had destroyed over USD 5 trillion in stock market values. However, only a fraction of the bubble was credit financed. Of course, banks had loan exposure to telecom, cable, and satellite operators, which suffered from the shakeout. But the resulting loan loss provisions were easily absorbed and never threatened the stability of the banking system. Unlevered private and institutional investors absorbed the losses of the dot.com bubble. The conclusion that asset bubbles were best left to sort themselves out on their own seemed obvious.

As global regulators were busy working on reforming the Basel I framework in the BIS towers in Basel, private sector banks worked simultaneously on the concept of economic capital (EC). The aim was to get a common denominator for risks. Whilst VaR captured risk in the trading books, economic capital attempted to capture the risk of the entire bank and define the amount of capital required to permanently staying solvent. As a concept, economic capital was widely popularised by JP Morgan's CreditMetrics. The statistical models underlying economic capital were an extension of VaR tools. However, the availability and accuracy of the data required for calculating EC was a challenge. Market risk data was easily available and mostly accurate. Credit and operational risk data however were neither. Both required precise measurement of event probabilities, the loss at any given event, and the recoverability of such losses. But there were so few observable events that they became statistically insignificant. Thus, many banks complemented their data sets with interpolations and estimates. The unreliability of the data became apparent in 2008 when losses, which economic capital models predicted to occur only every 10,000 years, materialised overnight. In the aftermath of the crisis, Alan Greenspan summarised it crisply: "We failed to comprehend the size of the expansion of so–called tail risk as became clear in the wake of the Lehman collapse. The tail was morbidly obese". Banks calculated economic capital numbers with a high degree of accuracy, but using flawed data made the results misleading and dangerous. It was garbage in, garbage out.

The availability of sophisticated information technology had seduced many banks to chase the one number, which would quantify a bank's risk profile. But overall, the development in risk management and corporate governance was slow. A few banks appointed chief risk officers on board levels but the systematic assessment of risks and rewards was not conducted. The quantification of risk with economic capital had created a false sense of security. Even worse, most board members did not understand the model they relied on. The wave of mergers, which dominated the financial industry in the late 1990 and early 2000, did not make matters easier. With widely varying IT systems, group-wide data aggregation became a real challenge. Most banks would head into the Great Financial Crisis with very limited and definitely incomplete management information about the risk on their balance sheet.

By 2004, the BIS finally published its Basel II recommendations. Its "advanced approach" resembled the economic capital in several aspects, relying on internal models and internal bank data, although the BIS had sharply reduced the diversification factors, which banks used in their own models. Nevertheless, for most banks, the new Basel II standards reduced capital for mortgages and corporate loans. Basel

II was eventually adopted in Europe and became law in 2007. However, America did not follow. Several regulators opposed its introduction. They objected to the reduction of capital in the system.

Whilst the regulatory world was busily focussed on credit risk, the big housing bubble in the USA was building up. Contrary to the dot.com crisis, this time the bubble was credit financed. Banks on both sides of the Atlantic had huge inventories of completely illiquid credit products on their balance sheet. They were mostly booked in their trading books and covered with a sliver of capital. Most of these inventories were short-term financed, pushing the leverage of the banking sector to a high level. There were no liquidity rules, which would limit the excessive use of short-term borrowings. The moment was missed in the run up to the Great Financial Crisis to adjust the regulatory framework to the changed business model of banks. When the Great Financial Crisis eventually hit the banking industry, it was a perfect storm. Banks did not have enough capital and liquidity. The management did not understand the complexity of its business model. The regulatory framework was weak.

15.9 A Rerun Avoided and Conclusions

The Global Financial Crisis wreaked havoc on the financial system and triggered the Great Recession destroying millions of jobs. The financial burden several European countries were forced to assume jeopardised the existence of the EURO and the coherence of the European Union. Without the active intervention of the Federal Reserve and the European Central Bank, both injecting trillions into the financial system, and without the rescue packages provided by governments around the world, a repetition of the Great Depression would well have been possible.

Many regulatory reforms have been enacted over the last 7 years to address the banks' shortcomings and to improve financial stability. Notably, the Basel process was significantly accelerated. The capital requirements for banks almost tripled. A new leverage ratio now limits the size of banks' balance sheets. Liquidity requirements found entry into the regulatory framework. The design flaws of the market risk rules were eventually addressed and, with regular stress tests, the resilience of the banking system strengthened.

However, it is important to recall that the weakness of the regulatory framework was only one of the factors contributing to 40 years of financial turbulence. The correct analysis and handling of the other four forces is equally essential to improve the stability of the financial system.

The first force was the re-establishment of international trade and free capital flows, which started with the arrival of petrodollars. Many of the financial troubles over the last 40 years have their roots in imprudent macroeconomic policies. A proper set of prudent macroeconomic policies is thus vital for the safeguarding of financial stability. In a world where institutional investors control more assets than banks and have to manage them for a return, investments will always flow from underperforming areas to more profitable opportunities. Hence, the integration of macroeconomic policies with prudential rules is vital. The Financial Stability Boards set up in Europe and the USA are a promising start. They combine central bankers, prudential regulators, and government representatives. The instruments are in place to take a holistic view. Alternatively, financial system could be organised strictly along national borders, but this would also end the free flow of capital and free trade, which create most of the today's wealth, jobs, and growth.

The second major force was information technology. From ATMs to electronic trading, it created many new business opportunities, which did not exist before and allowed banks to become much more efficient. It was inaccurate data, which misled banks and regulators about the inherent risk of their new business ventures, not technology. The continued advance in data management should enable the financial industry to build more reliable risk management tools. Today's technology is also more powerful than 20 years ago. Best-in-class banks now have digitised their balance sheet and train their managers like pilots in a flight simulator. But information technology will not lose its disruptive nature—it will continue to replace established business models and may well have the capacity to entirely disintermediate banks. Peer-to-peer lending platforms are already well established. Fund management platforms, which could render private bankers obsolete, manage already billions. The risk resulting from disruptive technology is to be taken seriously. Financial products are digital. Were it not for the complex set of regulations, the financial industry may well have had its Kodak moment already.

The third of the five big trends was the emergence of institutional investors. Requiring joint-stock companies to be profitable was a true game changer. Market discipline imposed by institutional investors was widely beneficial. It forced the manufacturing industry to leave or restructure unprofitable businesses and venture into new growth areas. Capital market discipline also forced banks to evolve. But in the absence of proper resolution tools to resolve banks without doing damage to an entire economy, the cost of failing banks fell into the lap of taxpayers. It is thus no surprise that in several European countries profitability targets for banks are still suspect. Resolving the clash between the public policy goal of financial stability and capital market discipline is thus a key priority. The bank resolution regimes adopted on both sides of the Atlantic as well as the recent proposals by the Financial Stability Board are all steps in the right direction. Institutional investors are here to stay. The need for our pension and insurance money to be invested will not go away.

Force no. 4 was the globalisation and ever-growing complexity of the financial system. As long as international trade and global capital markets exist, global banks will be needed. However, the case for financial conglomerates still needs to be made. Synergies are often named as to why retail and investment banking should be combined. Indeed, funding an investment bank with retail deposits is synergetic and specifically useful in times of liquidity stress. However, it also directly transmits investment banking risk to retail operations. This is not desirable. The recent discussion in Europe on ways to ring-fence retail deposits reflects this concern. In the absence of funding, there are few discernible synergies, which are not leading to conflict of interest or are plain cross-subsidies. It is also questionable, whether there is a need for the current complexity of banks' organisational structure with its thousands of subsidiaries. Often, these structures serve tax optimisation strategies or regulatory arbitrage. It may well be possible that market forces resolve the issue. Many experienced traders and M&A bankers have already left larger banks to set up their own funds or boutiques. As a prominent buy-side equity analyst from Goldman Sachs recently wrote, “the jury for big, integrated banks is out”. The regulatory capital surcharge for systemically important banks (SIB) may make the business model of integrated banks unprofitable. Splitting banks would also address an often-mentioned concern by investors that complex banks are difficult to analyse. Whilst the disclosure initiative of the Financial Stability Board made progress and improved disclosures, annual reports still count 500 pages and are a challenge to read.

Making the financial system less complex is not a goal, which relates to banks alone. Disentangling the financial infrastructure from banks would be a major step in improving financial stability as well. By making the clearing of OTC derivatives mandatory, a major step was made. The default of a major derivative trading house is not able to jeopardise the existence of the derivatives market again. Payment and security settlement systems should be set up in similar fashion. They could still be owned by banks but would have to be ring-fenced and bankruptcy remote. Given the fragmented liquidity of most traded financial instruments, dark pools of liquidity should be brought into the open and securities trading on exchanges be mandatory again. Competition is not the only tool to make sure that the monopoly of exchanges is not abused. The utility sector proves that appropriate supervision by government agencies can achieve the same goal.

Seven years ago, the Great Financial Crisis revealed with unmasked brutality how instable the financial system had become after the end of the Bretton Woods. The history of the last 40 years illustrates how important it is to adjust and redesign governance and regulations when the business environment fundamentally changes. Making timely adjustments will be critical for preserving financial stability going forward. The world will continue to change and technology will continue to be disruptive. The end of Bretton Woods did not cause the Great Financial Crisis. It was the inability of bankers and supervisors to adjust.

Chapter 16
Risk and Representation: The Limits of Risk Management

Heinz Zimmermann

16.1 Introduction[1]

Much was debated in the past years about the causes of the financial crisis which was triggered by the collapse of the US real estate market and the implied huge losses in complex-structured credit securities by large financial institutions, mostly banks. The crisis also reveals fundamental failures in the measurement, management, and transfer of risk in the financial system as well as methodological weaknesses (to say the least) in the regulation of financial institutions. Much has been learned about the (il-)liquidity of markets and its self-reinforcing effects, but the real problem is deeper.

I postulate that there was an insufficient awareness of the role of the representation of risks (such accounting standards, risk models, or management processes) with respect to the emergence of risks, in particular systemic risks and financial crises. The wording used in public commentaries is revealing and serves as an example: In the first months of the crisis, when banks announced their first substantial write-offs on their US mortgage-based assets, the "fear of further write-downs"

[1]The topic discussed in this *chapter* is partially beyond the scope of a financial economist. Nevertheless I hope that the thoughts are stimulating. I apologize for any shortcomings and errors and bear sole responsibility. The chapter is partly based on an earlier paper of the author, written in German in early 2008, shortly after the breakout of the financial crisis, entitled "Risiken und Repräsentation. Über Krisen des Finanzsystems" (in B. Strebel-Aerni (Ed.): Standards für nachhaltige Finanzmärkte, Schulthess (2008).

H. Zimmermann (✉)
Faculty of Business and Economics, Department of Finance, University of Basel,
Peter Merian-Weg 6, Basel 4002, Switzerland
e-mail: heinz.zimmermann@unibas.ch

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_16

was mostly regarded as the main cause of subsequent losses suffered by the banks' shareholders, not primarily the fear of actual further losses on the banks' risky positions!

It seems that if the object of interest is sufficiently abstract or complex (or both), as in the case of risk, the specific representation is a key determinant for shaping the perception or construction of reality: Credit rating procedures, accounting rules, write-downs, risk weighting schemes, regulatory capital standards, or monetary actions play a crucial role in the process of shaping the perceived risk of the financial system, and for rationalizing the potential losses to which financial intermediaries and their claimholders, or taxpayers, are exposed.[2]

Therefore, when analyzing the events in the progression of a crisis, it is notoriously difficult to discriminate between the "real" cause(s) of the problems and their multiple, partly self-reinforcing consequences. Fair-value-based accounting rules, rating, and model-based capital standards: did they passively reveal the "true" inherent risks of the system or did they cause, or at least accelerate, the risks? Did the various supervisory measures, monetary actions, or public financial stability programs mitigate or aggravate the events? How did all that affect the behavior of the economic agents?

An analysis of these questions should rely on a framework which is rich enough to represent the financial system as a complex, dynamically evolving system. More importantly, the *perception* and *management* of risks crucially depend on an adequate (viable) representation of the financial system. A simple model of unidirectional causality (usually in the form of a linear factor model where the causality runs from exogenous risk factors to position values) may be perfectly fine in normal times, but it provides an unreliable representation of the system in a situation of stress.

But if already the knowledge reflected in the representation is inherently incomplete, how should potential actions (regulation or risk management) be designed and evaluated? Suppose a network model is regarded as adequate representation of the financial sector or a part of it—how should we regulate a network? Our everyday systems are unable in mastering traffic jams; how can we expect to get a conceptual grip for a financial crisis? The suggested measures (for example, raising minimum capital requirements for banks) are usually based on a very simple, mechanical understanding of the functioning of the financial system. But the economic agents are innovative, constantly forming new expectations and adapting their ideas and models to the new conditions of the system. Therefore, with any intervention in the system, its reactions change as well. It is well known from Heinz von Foerster's

[2] It should be noticed that the expected losses as reflected in the global financial sector potential writedowns were constantly revised upwards between 2007 and 2009; starting at some US$400 billion, the estimates ended up at some US$4000 billion in April 2009, including all banks in the USA, Europe, and Japan, and including loans as well as securities (source: International Monetary Fund, *Global Financial Stability Report*, April 2009, p. 35). A different question is how these estimates are related to the actual losses of the banking (respectively, the entire financial) system, and how "losses" are defined.

model of nontrivial machines that such a system, although deterministic in nature, is analytically undeterminable, and the reactions seem random and unpredictable.[3] But would we consider this model, which is after all a standard representation of complex systems, a useful or rather useless basis for regulating the financial system? Or are we better off by choosing a simpler representation?

A lot has been written in the recent years about the triggering moments of the current financial crisis, the weaknesses of the financial architecture, the contagion effects of the crisis on the real economy, and the effects of the stabilization measures.[4] This *chapter* begins with a few general remarks about the representation of risk from an epistemological point of view (Sect. 16.2). The specific problems regarding liquidity risk are addressed next (Sects. 16.3 and 16.4), followed by a discussion of second-order knowledge traps and their circular effects, originating from incomplete or overrated knowledge (Sect. 16.5). From this background, the significance of standards in the financial system is discussed (Sect. 16.6). The *chapter* concludes with some final remarks (Sect. 16.7).

16.2 Risk and Its Representation

What is risk?[5] When talking to an option trader, "implied volatility," "the vola smile or smirk," is all that matters: It *is* the relevant uncertainty within his or her field of activity, experience, and perception. If the stock market crashes and the implied volatility triples, he or she would not have the slightest doubt that the uncertainty in the market has risen substantially. In technical terms, his or her perception or judgment of risk is fundamentally framed by the underlying pricing model[6]: in the simple case of the Black-Scholes model, fluctuations in the economic environment must be inevitably attributed to fluctuations in "volatility"—it is the only parameter in the model which is able to capture unobservable pricing factors. An observer with a perception that is not framed by the Black-Scholes model would possibly not

[3] Nontrivial machines are characterized by a state-dependent operating system (program), where the state is determined by the input to the system. Depending on the circularity of the system's architecture, the model can be used to analyze the dynamic behavior of systems such as learning, memory, adaptive behavior, and randomness. But the model also clarifies the limited knowledge which can be retrieved from the observed in- and outputs of a system about its internal unobservable "program"—even under very simple assumptions. See v. Foerster (2003), pp. 309–313.

[4] Among the numerous and almost uncountable references, the following are particularly worth reading from a financial economist's point of view: French et al. (2010) and Acharya et al. (2009).

[5] Although widely debated, in order to simplify the discussion, we do not distinguish between risk and uncertainty in this *chapter*.

[6] Notice that an implied volatility is only defined with respect to a specific option pricing model, which in turn depends on a specific stochastic process of the underlying securities and the implied arbitrage mechanism. In the Black-Scholes model, this reduces to an assumption about the standard deviation (volatility) of logarithmic price changes.

associate the market turmoil at all to "volatility" (a purely statistical representation), but to increased demand for protection, changing risk premiums, illiquidity, impaired market confidence, technical correction, and so on.

The option pricing framework with its "implied volatility" paradigm is an instructive example how models shape the perception and judgment of individuals in their professional work.[7] Notice that the claim here is not merely that the model serves as a complexity-reducing device (which it is indeed), but defines a linguistic code. It enables a standardized, fast communication in the hectic marketplace and the settlement of economic transactions. The pioneering work by Piaget, v. Glasersfeld and others demonstrates how language (in the broad sense, including e.g. mimic expressions) affects thinking and shapes, organizes, and structures people's perception. Language therefore determines what observers observe, what they construct, what they know and accept as personal reality. Therefore, as noted previously, the "implied volatility" *is* the reality, the (possibly hidden) model *is* the reality, and the "map *is* the territory" as v. Foerster bluntly stated.[8]

Although this view seems radical and controversial, it brings a novel perspective into the modeling of financial risk which stands in fresh contrast to the common "quest for the right model" mostly encountered in the theory of finance and financial management, and nourished by the administrative validation procedure in recognizing internal risk models under Basel II. Of course, the constructed "reality" may be regarded as being wrong and the individual may fail with the used model—but this is not the adequate perspective of the cognitive process under the constructivist approach: the emphasis is shifted from the "reality" to be discovered and re-presented by an adequate model to the creation (construction) of knowledge and the formation of perception. As discussed below, this seems to be quite a fruitful—viable—approach for our discussion. The reason is that risk and randomness are abstract, not directly observable phenomena of daily life. We only perceive realized damages, losses, accidents, crises, and the like. Abstract categories require an adequate cognitive (or mental) representation.[9] Risk is traditionally represented in many different ways: as narratives such

[7] Within the social sciences literature, a similar argument (with respect to option pricing models) can be found in MacKenzie (2006), although his focus is slightly different from the constructivist perspective advanced in the text here. A general analysis of the relationship between mathematical models and realities (ontological, personal, social, and formal) can be found in Henning (2009).

[8] See Foerster/Pörksen (1998), p. 82.

[9] The term "representation" is always tricky to use in a constructivist setting. Of course, representation does not only apply to abstract categories which are not directly accessible to sensual perception, but also refers to a general epistemological category. In this chapter, the term is used to characterize a device, by which objects—whether "real" (in the traditional meaning of an objective ontological reality) or "constructed"—are made accessible to our practical experience. While our point of view is fundamentally constructivist, we do not go as far as E. von Glasersfeld's who, in the absence of an observer-independent world in itself, abandons the expression right away (e.g., in v. Glasersfeld 1996). We agree with von Foerster that the term "re-presentation" (in German: "Ab-Abbildung") is inadequate or misleading with respect to a "reality" being "presented" (see v. Foerster 2013). However, our understanding is that representation reflects a state of knowledge (or a basis for acquiring knowledge) of an observing individual.

as myths and saga, as games of chance or lotteries, and most prominently as probabilistic and statistical models. Chance even found its systematic way in performing arts and the literature.[10] In business risk management, when using checklists or early warning systems, it is interesting to notice that the representation of risk is blurred. This insight is related to a general observation by sociologist U. Beck who claims that, in the process of defining (in our terminology: constructing) risk, "[t]he dimensions of the hazard are limited from the very beginning to technical manageability."[11] In the case of financial risk management, statistical models have a long and successful history, particularly in the field of insurance, as long as large ensembles of events can be represented by sufficiently stable statistical laws (damages, mortality, etc.).

Things became more difficult in the emergence of capital market risks (interest rates, commodity and stock prices): First, they can only be incompletely diversified which implies that comovement and temporal variability of prices are an essential part of modeling. Second, with the growth of option markets, the modeling of nonlinear, possibly even path-dependent risk profiles started challenging researchers. And third, with the rapid growth of over-the-counter (OTC) markets, counterparty (credit) and liquidity risk became issues of major public concern, not only since the financial crisis.

In technical fields (IT, engineering, or architecture), accounting, or international law, "standards" have a long tradition as quality, and hence risk management devices. It is therefore not surprising that the quest for generally acceptable financial risk "standards" has emerged over the past two decades, triggered by the banks' progress in implementing quantitative risk management tools and the regulator's recognition of model-based capital standards.[12] The point is discussed below (Sect. 16.6), but it should be obvious that standards are not primarily devices to represent risk, but to recognize or even approve the representation and management of risk. Standards should therefore be regarded as second-order representations of risk.

A final aspect is important: Under the constructivist, or pragmatic, epistemological view, the objective of scientific knowledge is not to discover (or represent) an ontologic truth or reality, detached from the observer's experience. Instead, knowledge is considered a "tool within the realm of experience" which is reflected in more or less useful inventions, fictions, or (in von Glasersfeld's

---

[10] A very illuminating collection of essays on this subject (in German, however) can be found in Gendola/Kamphusmann (1999).

[11] Beck (1992), p. 29; the original version was published in German (Beck 1986). Interestingly, in the English version, the quoted sentence is supplemented by the following remark: "In some circles it is said that risks which are not yet technically manageable do not exist—at least not in scientific calculation or jurisdictional judgment. These uncalculable threats add up to an unknown residual risk which becomes the industrial endowment for everyone everywhere" (p. 29). This is extremely revealing in the context of the systemic relevance of liquidity risk discussed below.

[12] At least, there are generally accepted risk management "principles" (GARP).

terminology) "viable" concepts or conceptual constructs such as actions, mental operations and structures, or theories.[13]

A model or business practice may be viable in one context but fail to be so in another. This sounds fairly trivial, but was largely overlooked in the modeling of risk. It implies that there is no horse race of models to find out which one represents the reality (a specific type of risk) in the "best" way. It has long been unrecognized that a model's objective in risk management, and thus its required profile, is usually not identical to other applications. A model, which is used to determine the value of derivative assets, should in the first place exclude arbitrage opportunities, and must therefore take the relevant market microstructure (tradability of hedge positions, short selling opportunities, transaction costs, taxes, etc.) into account. In the case of American options, for example, it is essential to take into account the optimal early exercise decisions in the valuation model. Should the same models be blindly used for the risk management of derivatives? This is indeed common practice, although the requirements on the entire model *architecture*, and not merely on the specified parameters, are different. The American option provides a good example: For the party who has a short position in the contract, the biggest risk does not arise from the rational (optimal) exercise behavior of the counterparty, but from a possibly irrational exercise decision, whatever the reason might be—the pricing model is, however, based on optimal behavior. Another example is the assumption of continuous hedging and replication possibilities, which may be a suitable approximation for pricing, but not for modeling the liquidity risk of a market. Using stress tests in this respect is short-sighted, because these work on the level of parameter specification and risk scenarios, but the assumptions of the model and the range represented risks are not questioned. In fact, stress tests should invalidate or mutate the models' architecture, which, for example, could be represented in the form of genetic algorithms in a more viable way.

It is not only crucial to constantly validate the objective of models, but also to be aware that models explicitly shape what we do and do *not* consider a risk within the range of a specific risk management task. Of course, we have been long aware of "model risk" in risk management, and perform many calculations and estimates using alternative assumptions from different models.

But the problem goes deeper, however, and arises from the self-referential character of the nature of risk, i.e., the risk of risk. We have to find a *risk-adequate* representation of risks, meaning a form of representation (similar to a language) of risks, which reveals the risk of representation itself and keeps the represented object "alive." The analogy with language is useful because the linguistic sign system is not only a basic example for the representation of knowledge, but also shares many structural properties, such as self-reference and circularity: Language can be used to

[13] As mentioned earlier, von Glasersfeld strictly avoids the term "re-presentation" in his characterization of cognitive action, because he relates it to an unknowable ontological reality. However, in a less strict view, the "concepts or conceptual constructs" also require a representation, a representation reflecting a certain state of knowledge. The quoted phrases are from v. Glasersfeld (1998).

talk about language (i.e., by inventing a meta-language).[14] Of course, the linguistic rules could be constructed in a way that the language is gradually distorted or destroyed in the circular process, in a way that the ability to describe itself is abandoned. This would completely invalidate the epistemological function of language and communication. Similarly, a system which is designed for representing risks and which, through its ongoing circularity, would make the risks of the represented objects gradually disappear would be as worthless and inoperable as a language which had no ability to describe itself.[15] The inclusion of such circularity in the design of descriptive systems such as risk management in itself represents a problem that is far from trivial, as discussed in the following sections.

16.3 Liquidity as an Information Problem

Liquidity risk, and the various types of risks assigned to this category,[16] is a particularly insidious risk category because of its self-destructive impact on the information processing capacity of markets. In extreme cases, this may mean that the market mechanism generates no information at all which disables the communicative function of markets and thereby invalidates the representation of market-based security values and their risks; in a weakened form, the information capacity due to high spreads and low conditioning volumes can be so severely restricted that the information for the assessment of risk positions is too blurred and thus unsuitable. In more extreme forms, the information may even be misleading. In the context of market-based risk and accounting procedures, this represents a "second-order" problem because, in a sense, the representation of risk makes the represented "object" disappear.[17] The object disappears, but not the risk itself, and this is where the dilemma is! Therefore, returning to the question of risk-appropriate representation of risks relating to illiquidity, we are confronted with a special, yet largely overlooked difficulty: which are the relevant objects to be represented in the presence of illiquid markets?

A further loop or self-reference in the information system originates from the use of model values in lieu of the missing market values; this practice is consistent with market-based accounting principles and is recognized by supervisory authorities. This generates, in a sense, a representation problem at a higher level. At the same time, most accounting standards require depreciation of positions which are becom-

[14] See v. Foerster (1997), p. 165 (original German edition).

[15] In his Tractatus, Ludwig Wittgenstein characterized the language as located on the very limit between the speakable and the unspeakable. I would suggest that representations of risk, e.g., by a probabilistic risk model, are similarly located at the borderline between safety and uncertainty.

[16] Two typical forms of liquidity risks are "market" liquidity and "funding" liquidity; their relationship is analyzed by Brunnermeier/Pedersen (2009).

[17] Or in attenuated form: the object (market price) which is required for the representation of risk loses its suitability or quality.

ing illiquid. Finally, financial analysts and the media need to provide information about the market valuation of the affected financial institutions. This is a nearly endless loop of self-references, which ultimately deprives the risk management system and the financial market of any information function: the system starts to produce noise rather than information.

Ultimately, it is no longer possible to distinguish between the risk of (inadequate) representation and the represented objects, and hence the representation of the risks. The often-raised question of how many secondary effects are generated “by the system itself,” i.e., are fictional in nature and not related to the financial system intrinsically, seems pointless in this context. The question has no answer, because it would require a natural break at some point in the operation of the system, from where we could distinguish between the “original” problems (real estate crisis, excess leverage of the system, etc.) and the “subsequent” effects. This view is incompatible with the modern view of dynamic systems characterized by circularity and feedback effects.

A particularly strong case of circularity is associated with the interpretation of credit risk (i.e., the quality of debtors and counterparties). The criticism directed at the rating agencies shows this clearly: Should rating agencies review credit ratings based on functioning market structures and economic conditions (i.e., “conditional” reviews), or does the market expect *un*conditional reviews? The high correlation between liquidity and credit risks[18] highlights the importance of this issue. In this context, the question of the temporal (i.e., causal) relationship between credit and liquidity risk is particularly relevant. Is credit risk determined by the (il) liquidity and market frictions, or do major credit events and related information problems trigger liquidity crises? The effect may well be circular and self-reinforcing.[19]

16.4 Liquidity as a Network Problem

Liquidity risk is multidimensional; it is related to many different aspects of risk. In the context of risk management or the financial regulation, liquidity is often regarded as a separate category of risk, but from an economic point of view, a representation as superposed or second-order risk category would be more adequate. In particular, an adequate representation should also account for the *systemic* nature of (il) liquidity.

It is well known that disruptions in market liquidity are not only followed, but often even caused by coordination problems between the market participants in

[18] Some early estimates about this relationship are reported in the *Financial Stability Review* of the ECB (European Central Bank), June 2008: the reported correlation coefficient for a cross section of 10 countries is roughly 0.8.

[19] An insightful analysis of the liquidity-credit risk spiral during the financial crisis can be found in Brunnermeier (2009).

times of substantial uncertainty and financial turmoil: lack of transparency about the size and nature of counterparties' risk positions, their intended risk behavior, trading intensions, or deficits in the informational role of the price system may have self-reinforcing effects.[20]

If coordination problems matter, a network model likely represents a viable representation for liquidity risks. The key merit of this approach is the shift in the perspective from the single player down to the architecture of the network, in particular to the rules affecting the coherence of the system. Also, the network approach puts into question the ability to attribute the dysfunction or vulnerability of a system operating as a network to a clearly definable cause and to derive simple, promising stabilizing rules. The following metaphor by cyberneticist H. von Foerster (2002), p. 133, provides an apt illustration of this point:

> *If I pull a crocheted vest, and I pull the thread in one place, the whole vest unravels suddenly. Now we can say: 'At the place where I pulled, there lies the essence of the vest'. And because it is located there, the whole vest disappears when I destroy this part. What ceases to be seen here is the network of threads, the 'vest network'. In a vest network, in which a thread is connected to the others, such dissolutions can emanate from a single point; such destruction can be initiated continuously from any point. Conversely, this does not mean that the pulling point would be the location of a certain function that is now no longer available and can be identified. The system does not work anymore because damage in one place can spread to the whole thing.*

That the financial crisis was triggered by the problems on the American real estate market or structural defects of mortgage securitization should not be taken, from this perspective, as an opportunity to align the efforts to stabilize the financial system too closely with these specific factors. This is because errors with similar or perhaps even greater systemic impact could have occurred in other places of the financial system. This approach is consistent with the interpretation of various experts and stability reports of monetary authorities which indicated the fragility of the financial system at an early stage, due to various structural changes,[21] but ultimately did not foresee the initial trigger moment of the crisis in the real estate market.[22]

---

[20] The literature analyzing the stock market crash of 1987 has emphasized many of these problems. The informational externalities related to invisible and uncoordinated dynamic portfolio insurance strategies and their effect on market liquidity was studied in detail by S. Grossman. A selection of his papers can be found in Grossman (1989).

[21] Concerns were related to structural weaknesses and the potential insolvency of central counterparties of credit derivatives (which were the major, highly leveraged investors of US mortgages), the strong dependency between hedge funds and investment banks, cross-border and cross-currency issues in the lending process between central banks, or dysfunctions in international clearing and settlement transactions. See, e.g., Zimmermann (2007) for a discussion of the subjects in the public concern just instances before the breakout of the financial crisis.

[22] Only a few eminent economists can be credited to having foreseen a financial crisis caused by structural deficits of the US real estate market. Nobel laureate Robert Shiller is a prominent exception. Also Frankel (2006) reveals the structural weakness of the US subprime market in great detail, without implying a global financial crisis however.

The analysis of the architecture of the financial system should remain the primary concern in the future. However, the financial crisis teaches us that, in the reform of the financial system, more efforts should be directed to information and coordination problems related to the illiquidity of (apparently unrelated) market segments, and to the impaired funding and payment capacity of major counterparties. However, the "systemic" view was ignored until recently in the regulatory discussion. Reports and research papers which addressed liquidity management and requirements of systemic relevant banks, even released after the breakout of the crisis,[23] did not consider systemic aspects. Liquidity management is regarded as a completely customizable management function free from any systemic considerations or consequences. But systemic problems require systemic solutions, or at least solutions that are directed at collective actions, as postulated by Eichberger/Summer (2005) in the context of banking regulation:

> *If regulation aims for risk allocation across the entire banking system, then it has to stop concentrating on individual bank balance sheets. (…) A systemic approach to banking regulation is just the beginning.*

This must be understood as a general device for the design of stabilizing measures for the financial system, and not just for the banks and their regulation.

To conclude, it should be emphasized that the solutions to the problems discussed in this section are not only found in systemic requirements (rules or standards for the processing of C&S transactions, a transparent market architecture for OTC transactions, etc.), but also involve rules of conduct applicable for individual institutions—as long as they are systemically meaningful. The requirement that every vehicle must have brake lights obviously represents a purely individual regulation, the need for which, however, fully results from a systemic requirement (because "I" do not need any brake lights for "my" vehicle). Therefore, rules with a systemic focus can be easily built in the traditional institutional based regulatory and supervisory framework; a good example are rules about securities' collateralization.

16.5 Limited Knowledge and Second-Order Knowledge Traps

What lessons can be drawn from the financial crisis? What are the implications of the preceding remarks? What exactly should be done for the financial system to better absorb such dysfunctions in the future? Is more capital needed for financial intermediaries, and how much?

As a matter of fact, knowledge about these issues is incomplete, and furthermore, even the knowledge *of* the knowledge is incomplete. Specifically, we know very little about what we would need to know to answer these questions. The financial

[23] See, for example, the consultation paper from the Basel Committee on "Principles for Sound Liquidity Risk Management and Supervision" (BIZ 2008).

system was not designed to be easily understood. Knowledge plays a major role for the design of modern financial markets, financial instruments, institutions, and processes—but it is not the same knowledge which is needed to understand the consequences from a complex process of financial innovation and its interaction with the financial system. Therefore, the systemically relevant knowledge is largely overrated in the financial sector.

Additional factors have contributed to an overestimation of knowledge in financial matters:

(a) The status of expertise

In the general perception, scientific competence and expertise are under control of academic researchers. In the past decades, consultants have taken over this function more and more. In contrast to the scientific experts of the past, the consultant charges high fees, and this can only be justified by delivering know-how and specific insights, i.e., by signaling competence. A side effect of this development is that this kind of knowledge is no longer publicly accessible and consequently no longer exposed to the scientific discourse. This applies of course to all areas of expertise developed outside of the scientific system, for example the research undertaken in banks, stock exchanges, and even by regulatory authorities.[24] The incurred risks for risk management should not be underestimated.

(b) The scientific process

The scientific process has an impact on the choice and nature of analyzed problems. Mathematics and statistics are essential and viable tools for the representation of risk in financial markets. But the quality of a tool, its internal logic and structure must be strictly distinguished from the constitutive properties of the object to be represented or constructed. Limited knowledge, indeterminism, or incompleteness are well accessible to the mathematical formalism,[25] but scientific "reductionism" directs mainstream research in more comfortable fields: v. Foerster contemplates on the method of inquiry employed by the hard sciences:[26]

> *If a system is too complex to be understood it is broken up into smaller pieces. If they, in turn, are still too complex, they are broken up into even smaller pieces, and so on, until the pieces are so small that at least one piece can be understood. The delightful feature of this process, the method of reduction, "reductionism", is that it inevitably leads to success.*

This implies Foerster's Theorem Number One: "The more profound the problem that is ignored, the greater are the chances for fame and success." The publication pressure, which is particularly powerful in scientific disciplines, thus led to a flood

[24] See Zimmermann (1999) for a detailed discussion of this development. The self-confidence, or arrogance, by which this knowledge is communicated to the world outside is sometimes remarkable. The wording of a white paper published by a major investment bank is revealing: "More than You Ever Wanted to Know about Volatility Swaps (But Less than Can Be Said)."

[25] By Gödel's theorem, incompleteness is even an inherent property of arithmetic systems.

[26] From "Responsibilities of Competence," in: v. Foerster (2003), pp 191–197. Originally published in 1972.

of technical results, which give the appearance of a huge knowledge, suitable for the purposes of risk management. But the formalism got increasingly detached from the nature of the problem to be analyzed, and the knowledge to be represented.

(c) Survivorship, success, and knowledge

In financial matters, it is often hard or even impossible to draw a clear distinction between signal and noise, or skill and luck, on typical statistical confidence levels. Still, success is commonly interpreted as a direct indicator of competence and knowledge. The selection process of success (*survivorship bias*) is ignored: in a system which accidentally distributes success and failure, the probability of success to survive naturally dominates the probability of failure. Self-selection makes the successful to survive, and makes them talk about their ongoing success. However, inference regarding superior knowledge is not justified. The selection process also leads complete randomness to appear like competence and knowledge to the public. If the selection process is connected to a high level of monetary compensation, people will be particularly tempted to attribute this to a high level of competence.

(d) The media and the quest for simplicity

The public overrates experts' knowledge because people have a strong desire for simple explanations of complex matters. The media eagerly respond to this desire and provide a steady stream of commentaries and statements from professionals. The search for eligible candidates is terminated if anyone is willing to give their opinions and explanations, and this process always ends successfully. Those remaining silent, because he knows that he knows little or nothing, are ignored.

In the public arena, from the point of view of the layman, expertise is also overrated because one needs knowledge to appreciate the value of knowledge. This creates a second-order effect of limited knowledge: how can ignorance be recognized and built in the architecture of observing systems? Risk management, in this respect, needs to find an adequate representation of knowledge about existing knowledge, and complementarily, knowledge of ignorance or at least the limitations of knowledge.

Notice that the argument is not about the deficits of possible representations of risk, e.g., what is typically known as "model risk" in the field of risk management. The claim is that the level of knowledge, and complementarily the level of ignorance, represented implicitly by a specific model ultimately amplifies the risks which ought to be represented.[27]

The idea can be further illustrated with an analogy to linguistics: nonrecognition of the non-expressible of a given vocabulary leads to a threat from *that* which cannot be expressed. A striking example of this dilemma is provided by Odysseus,

[27] This may sound rather abstract. An example: Metallgesellschaft in the early 1990s sold long-term commitments to deliver gasoline and hedged the exposure thereof by rolling over short-term futures contracts. The company assumed that futures markets remain in backwardation. When the futures market turned into contango, the company was forced to adjust the hedging strategy. Because of the substantial market share of the company and the illiquidity of the market, this adjustment amplified the adverse price behavior, i.e., increased exactly that risk which ought to be hedged. The company finally got insolvent. A detailed analysis can be found in Culp/Miller (1995).

who, upon entering the cave of Polyphemus the Cyclops, cunningly presents himself as “Nobody.” This benefited him after he rammed a stake in the eye of the sleeping giant who was holding him as a prisoner. With his cry of “Nobody has blinded me, Nobody has tried to kill me” Polyphemus could not secure the help of the other Cyclopes. If Polyphemus had recognized the limitation, or ambiguity of his vocabulary as imposed by Odysseus, his fate would have taken a better turn.

We conclude the preceding thoughts with the metaphor of eyeglasses: as a matter of fact you need glasses to realize that you actually need a pair of glasses. Without, you cannot even come to recognize that there is another, better representation of reality, one that is associated with fewer or lesser risks. The epistemological conclusion is that you can’t see what you can’t see. The double negation, however, does not mean that you can see! The implication (advanced by Spencer Brown, Heinz von Foerster, or Gotthard Günther) is that with self-referential processes classical logic fails by violating the principle of double negation. This has crucial implications for the design of risk management systems, albeit ones which are barely discussed. Returning to the metaphor of the eyeglasses, the circularity and the resulting dilemma for the risk management are obvious: how can you find a lost pair of glasses without glasses?

This is not an impossible task—provided that *you know the risk* of misplacing the glasses (i.e., you know the risk of not seeing). But the possible solutions have nothing to do with traditional actions, such as correcting the lenses. Rather they could involve, for example, an adaptation of structures (e.g., keeping the eyeglasses around your neck), the introduction of standards (e.g., strict rules about where the risky object should be deposited), or the implementation of intelligent search processes (e.g., pressing an emergency button to ask someone for help). The final section deals with some thoughts about this topic.

16.6 Standards

What is the role of standards for the representation of risk? In 2009, the International Organization for Standardization released its Risk Management Standards (ISO 31000, Principles and Guidelines) which should help firms to improve the quality of their processes in terms of “economic performance and professional reputation, as well as environmental, safety and societal outcomes.” The principles as well as the recent technical follow-up report (ISO/TR 31004:2013) are designed for general organizations, not specific sectors or activities. There are no specific standards for financial institutions, but general principles or best practices.[28] But obviously, international standards play an increasingly important role in international financial

[28] The use of “standard” is somehow ambiguous in the literature. In the field of law and prudential regulation, the term has a rather broad meaning (e.g., includes the principles released by the Bank of International Settlements). A decade ago, Nobel (2005) lists more than 60 standards in use in the field of international financial regulation.

operations (e.g., IT, clearing and settlement, law, accounting, audit, controlling) and shape the understanding and perception of risk.

The benefit of standards originates from the combination of commitment and flexibility. Typically, a standard defines a minimum level of quality (or regulation) that allows the individual jurisdictions enough leeway to agree upon a set of rules for specific needs. From a polito-economic point of view, a standard is usually the only way to establish rules at the international level.

Standards can be regarded as a prerequisite to the solution of several of the previously discussed issues, for example:

- Standardized approach in the representation, processing, and interpretation of information and data, and thus improved communication and coordination of decisions (e.g., accounting principles, trading statistics)
- Common terminology, definitions, and language (e.g., classification of products, risk categories, etc.)
- Reduction of technological frictions, operational inefficiencies, and therefore improved coordination of complex processes (e.g. in securities trading and settlement, payment systems, etc.)
- Mitagion of legal uncertainty by regulatory or contractual standards (e.g. netting rules, capital adequacy, collateralization)

Thus, standards seem to be promising for reducing the risks caused by various types of frictions and their implied second-order effects, i.e., the risks originating from an inadequate representation of risk. However, their overall performance may be difficult to assess.

This may be particularly true with respect to the aggregation of information. When defining a technological standard, e.g., the specification of securities numbers, less emphasis is put on information and knowledge aggregation issues than, for example, in the release of clearing requirements of OTC derivatives or capital standards of banks. Here, the administrative negotiation process involves procedures and activities (e.g., advice from experts, research, practical experience) so that the resulting standard signals new knowledge and leads to positive information and incentive effects. However, precisely the Basel II "bank capital standard" has been criticized for having destabilized the financial system and failed to prevent or mitigate the financial crisis.[29] The case illustrates that the process of establishing and implementing standards is not immune to administrative momentum and strategic private interests in negotiation. In light of the previous observations about the limits of knowledge, and knowledge of these limits, it should be considered that

- Standards determine the perceptions of the agents, both on a personal level (cognitively, psychologically) and an institutional level (e.g., through the legal and accounting framework).

[29] Admati/Hellwig (2012) provide an in-depth discussion of financial regulation, the role of banks' capital, and the safety of the banking system.

- Standards signal superior knowledge; they might negatively affect agents' incentive to process information and to acquire knowledge, or to manage or mitigate risks.
- Standards create incentives for opportunistic behavior by delegating responsibility to the standard-setting instance or authority; that is, standards are not considered as minimum quality requirements, but as defining the maximum required effort.[30]
- Standards cause, with high probability, a synchronization of the perception and behavior of agents, which decreases the heterogeneity of decisions and thereby damages the liquidity of the financial system as market coordination mechanism.

It is therefore essential that standards in the field of financial risk management are not narrowly focused on operative, or technical, matters but generate sufficient incentive for a *risk-adequate* representation of risks—i.e., forms of representation which reveal the risk of representation and keep the represented object "alive." Too far-reaching standards, however, which affect the behavior of heterogeneous agents in a unidirectional way, have the potential of having counterproductive effects by limiting or impairing market liquidity with the adverse effects discussed before.

16.7 Concluding Remarks

The management of risk in financial institutions has long been dominated by probabilistic and statistical models. This is not wrong, but incomplete. The recent financial crisis has again highlighted the limitations of a formal framework for the representation of risks, particularly if second-order effects associated with the representation of risk (risk of risk) are ignored. Specifically, the modeling and regulatory treatment of financial risks are blamed as amplifying factors of the crisis itself.

The discussion of this chapter focused on those risks which emerge from limited market liquidity. It provides a perfect case where there is a risk in the representation of risks: For example, the Basel II capital standard was conceived in view of perfectly operating markets whose associated risks can be represented by market price fluctuations and a sophisticated system of credit qualities standardized and formalized by rating agencies. Furthermore, the residual category of operational risks appeared to cover all other risks, but what about liquidity risks and how should they be treated? The unsatisfactory representation of this category is not only a symptom, but most likely the cause of some of today's problems.

[30] A good example is the capital ratio of banks which has consistently decreased over time with the implementation of stricter and more sophisticated *minimum* capital requirements. The increasing cost of equity capital is consistently used to rationalize this trend by the banking industry. Of course, causality runs in the opposite direction—whether it is a moral hazard issue or perception bias remains open here, but the effect on the systemic risk is striking.

[31] One of the editors of this volume, Robert A. Schwartz, can be credited as one of the pioneers in this field.

In addition, until the financial crisis, there was a pronounced research deficit in financial economics for modeling and understanding liquidity as a macroeconomic phenomenon. Market microstructure theory provides an analytically and empirically rich foundation for understanding the functioning of the financial markets as institutions, practically as microorganisms: the influence of the stock exchange architecture on the price discovery process, the behavior of market makers, or the determinants of bid-ask spreads.[31] Monetary economics, on the other hand, deals with issues such as optimal control of Central Bank liquidity, liquidity requirements of banks, and their macroeconomic effects. Finally, microeconomics is interested in the behavior of individuals facing incomplete information during banking crises (*bank runs*) or panic on securities exchanges (*herding*), and analyzes the influence of regulatory interventions. But these individual elements offer no satisfying picture of the financial system as a whole, because they do not fully represent its complexity. But complexity is a major constitutive feature of the modern financial system!

Although the convergence between monetary and financial economics progressed over the past years, in particular since the financial crisis,[32] more efforts are needed particularly on the methodological side. Strong focus should be placed on the risks of prudential regulation and risk management practices—meaning the systemic role of models, standards, management processes, algorithms, etc.—and should therefore be concerned with the representation of second-order risks and their circularity.

This is less complicated than it appears at first glance, because methodologically similar issues are being intensively worked on in other disciplines. It would be interesting, for example, to take advantage of the possibilities of computer science or computational economics. In the first section of this chapter, risk management was characterized as a circular, self-observing, and self-constructive system to represent risks. A methodology developed at the computer science department at the University of Basel uses exactly the circularity of processes to improve on the robustness of processes, specifically in the case of a programming language faced with external disturbances.[33] The circularity and self-reference become manifest in the program creating a code which contains its own description. This circularity in the form of self-replication allows the program to repair itself when disturbances occur; this error-correction mechanism significantly improves the stability of the system. Self-replication and self-repair are, after all, constitutive features of biological systems and were mathematically studied long ago (e.g., by John von Neumann and others). It would be interesting to investigate the possibilities—and *failure*—of these systems for the representation of risk, and second-order risk, in the financial sector along the lines of this chapter.

[32] The work by Nobel laureate Jean Tirole, and more recently by Markus Brunnermeier (see, e.g., the referenced article) and Hyun Song Shin, to mention just a few representative researchers, is remarkable. They focus on the relationship between market liquidity and financing patterns (leverage, collateralization) to analyze contagion effects and financial stability. Shin (2010) gives an overview on this research. Shin's nomination as Economic Advisor and Head of Economic Research at the Bank of International Settlements (2014) is a promising perspective in this light.

[33] See Meyer/Tschudin (2012).

Chapter 17
T2S: Creating a New Post-trade Landscape

Karla Amend and Matthias Papenfuß

Today, although being highly efficient in national markets, the European settlement infrastructure provides a complex, highly fragmented picture, which is characterised by national specificities and barriers. This prevents competition and efficient cross-border settlement at reasonable costs and also comes with operational risks. Target2-Securities (T2S) sets out to provide an infrastructure which will remove those handicaps. But once the infrastructure is in production starting in 2015, T2S is supposed to act as a game changer which will initiate a sustainable transformation of the whole post-trade securities servicing industry.

At the end of the 1990s the creation of a single EU market had top priority on the European political agenda. The financial service industry was expected to play an important role in order to achieve the free movement of goods, services and capital. At that point in time, several initiatives have been initiated around trading, clearing and settlement that aimed at analysing the current status in the European countries and at evaluating appropriate measures to achieve the overarching goal of a single EU market.

Amongst the various initiatives the work of the Giovannini Group had put the spot on the clearing and settlement layer with a focus on the specific deficiencies around cross-border clearing and settlement. The group issued its first report[1] in 2001 in which 15 main barriers have been identified for causing the existing problems. The barriers mainly stem from national market practices, regulatory requirements, tax procedures and legal uncertainty. Accordingly, both the public and the private sectors had initiated activities to remove the barriers that cause the fragmentation of securities settlement, which takes place along national borders and in more than 30 different systems within the EU. In order to cope with the different conditions on a legal, fiscal,

[1] The Giovannini Group "Cross-Border Clearing and Settlement Arrangements in the European Union," November 2002

K. Amend (✉) • M. Papenfuß
Clearstream International, Mergenthalerallee 61, Eschborn 65760, Germany
e-mail: karla.amend@clearstream.com; mathias.papenfuß@clearstream.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_17

operational and technical level, market participants have to involve a significant number of intermediaries such as agents and custodians. The 2001 Giovannini report revealed that a common cross-border equity transaction would require the involvement of at least 11 intermediaries (compared with only 5 for an equivalent domestic transaction) and a minimum of 14 instructions per trade between parties. This causes not only tremendous costs for the securities service industry but also operational risks. In addition, you may add to the bill opportunity costs that result from cross-border activities not taking place due to complexity and costs. In essence, the situation is not compatible at all with the requirements of a single financial market and below the standards achieved in CeBM settlement.

In their second report,[2] the group presented a strategy for removing the identified 15 barriers. From the perspective of integration, the priority barriers are those imposing restrictions on the choice of settlement location activities. Removing these barriers will enable investors to choose the location for their post-trading activities and thus trigger a market-led integration of clearing and settlement arrangements across the EU. A significant increase in cross-border securities trade is to be expected but unacceptable levels of operational and legal risk will still persist, if other barriers remain in place.

In June 2008, the ECONOMIC and FINANCIAL AFFAIRS Council (ECOFIN) provided its conclusions on clearing and settlement,[3] which clearly encouraged the European Central Bank (ECB) to provide as soon as possible to interested CSDs an offer for T2S. ECOFIN also emphasized the need to trim down the projected end-to-end costs and to achieve a reliable business case for T2S. Based on the ECOFIN conclusions, the positive experience made by the introduction of TARGET2 (the real-time gross settlement system for EUR cash payments in central bank money, operated by ECB), the substantial support received from several market consultations, and considering the potential efficiency increase that could derive from holding both securities and CeBM accounts in an integrated technical environment, still in 2008 the Eurosystem decided to develop T2S for providing security settlement services to CSDs. The benefits associated with the integration of the European securities infrastructure were expected to outweigh the challenges of T2S.

Main benefits: T2S will provide the technical infrastructure required by a single market and currency, enabling participating CSDs to offer DvP (deliver vs. payment) settlement of securities transactions in CeBM one technical platform, with harmonised operating times and deadlines, operational rules and communication messages under a common legal framework. This ultimately means that domestic settlement in Europe becomes harmonised. Furthermore, in terms of cost, risk and technical processing cross-border settlement becomes identical to domestic settlement. The harmonisation of standards and market practice will reduce complexity in the post-trade layer of the European securities markets and thus make

[2] The Giovannini Group "Second Report on EU Clearing and Settlement Arrangements," April 2003.

[3] 2872nd ECONOMIC and FINANCIAL AFFAIRS Council meeting, Luxembourg, 3 June 2008; http://www.eu2008.si/en/News_and_Documents/Council_Conclusions/June/0206_ECOFIN.pdf.

securities settlement safer and more efficient while increasing competition and business opportunities. Also, it is expected that T2S will act as a catalyst for further harmonisation. Users of the new T2S platform will be able to settle securities in multiple CSDs from a single cash account in CeBM, and to move securities more easily and quickly across borders to where they are needed for collateralisation (and other) purposes. As a result, the market expects liquidity and collateral savings to be among the main benefits generated by T2S. Several market-driven studies aimed to quantify these benefits. According to a PWC study[4] banks will be able to reduce their Tier 1 capital needs by EUR33bn which is about 11 % of the total capital requirement of €295 billion projected for the Eurozone. A further study commissioned and conducted by Oliver Wyman[5] projected significant capital, funding and operating cost savings by delayering and consolidating their securities and cash holdings. The estimated annual benefits of three case studies range between €30 million and €70 million.

But before enjoying the benefits, the post-trade industry is now facing the largest infrastructure project this industry has ever seen. The market participants will have to assess the T2S impact on their business models which goes beyond pure adaptation efforts. T2S will initiate a reshaping which may range from simple adaptations of service levels up to a comprehensive business model transformation—depending on the participants' current position and their appetite for an extension of their existing service portfolio. In addition, the industry will have to swallow the estimated development costs of about EUR 1 billion and the project risks associated with such a long-term large-scale infrastructure project with multiple stakeholders.

From 30 CSDs that initially have shown interest in becoming T2S participant by signing the T2S Memorandum of Understanding in 2009 and 2010, thus far, 24 European CSDs (thereof 19 CSDs based in the Euro area and 5 from non-Euro-currency countries) out of 21 European markets have entered into the contractual agreement with the Eurosystem and will outsource their securities accounts to T2S for settlement purposes. The participating CSDs account for almost 100 % of Euro volumes currently settling in the Euro zone (Fig. 17.1).

On the cash side, 19 central banks will open dedicated cash accounts in Euro for their participants in T2S, so that settlement of securities against CeBM can take place in an integrated manner. The coverage of T2S may further grow in future, as other European central banks (and possibly currencies) and CSDs may decide to join the platform.

Nevertheless, an integrated platform for securities settlement as such will not deliver a fully integrated market on its own. Surrounding activities of EU authorities and the industry will contribute to increase the level of harmonisation in order to maximise the efficiency of cross-border settlement in T2S. Hence, it will ensure market access and equal conditions for all participants.

---

[4] PricewaterhouseCoopers "The 300-billion-euro Question—Survey on the Benefits of Target2-Securities," August 2013.

[5] Oliver Wyman "The T2S Opportunity—Unlocking the hidden benefits of Target2-Securities," September 2014.

CSD	Market Share
CBF	38,73 %
Iberclear	17,90 %
Euroclear France	15,43 %
Monte Titoli	13,28 %
Euroclear Finland	3,91 %
Euroclear NEGICEF	3,38 %
VP Securities	2,35 %
Euroclear Belgium	1,21 %
OeKB	0,79 %
Keler Hungary	0,69 %
Interbolsa	0,62 %
CSD of Romania	0,61 %
NBB-SSS	0,35 %
SIX SIS	0,33 %
BOGS	0,10 %
CSD Slovakia	0,09 %
Lithuania CSD	0,06 %
KDD Slovenia	0,06 %
EVK – Estonia	0,05 %
Cyprus Stock Exchange	0,03 %
Malta Stock Exchange	0,01 %
VPLux	0,00 %
LuxCSD S.A.	0,00 %
BNY Mellon CSD	0,00 %
Latvian CSD	0,00 %

Fig. 17.1 Overview of participating CSDs and their market share. Source figures (2013): http://sdw.ecb.europa.eu

The delivery responsibility of T2S has been given to four central banks (Deutsche Bundesbank, Banque de France, Banca d'Italia and Banco d'Espagna) which have been tasked to create a single platform to which all CSDs would outsource their settlement and accounting activities. This platform, T2S, will then be operated by the Eurosystem/ECB. T2S unites the securities part and the cash part of a settlement vs. payment (DvP) instruction onto one single platform in CeBM. Besides harmonisation of securities settlement processes, T2S is also tearing down the difference between settlement of a domestic and a cross-border transaction. As a result, the process of settling a (e.g.) French security between two (e.g.) French parties becomes exactly the same as settlement of the very same security between a (e.g.) German and an (e.g.) Italian party. All Euro-currency markets are in terms of securities settlement treated like one domestic market (Fig. 17.2).

Each CSD has signed the Framework Agreement with the ECB, which regulates the outsourcing of the handling of instructions (settlements) and balances (accounts) to T2S. The "golden source" for instructions and balances is kept in T2S, while the CSDs remain legally responsible for the data. In turn, the CSDs are holding the "golden source" for all client and securities (e.g. ISIN) related reference data and provide a copy of the relevant data to T2S. Services such as corporate actions/asset services, lending and borrowing, collateral management as well as connectivity from and to T2S remain under the full control and responsibility of the CSDs.

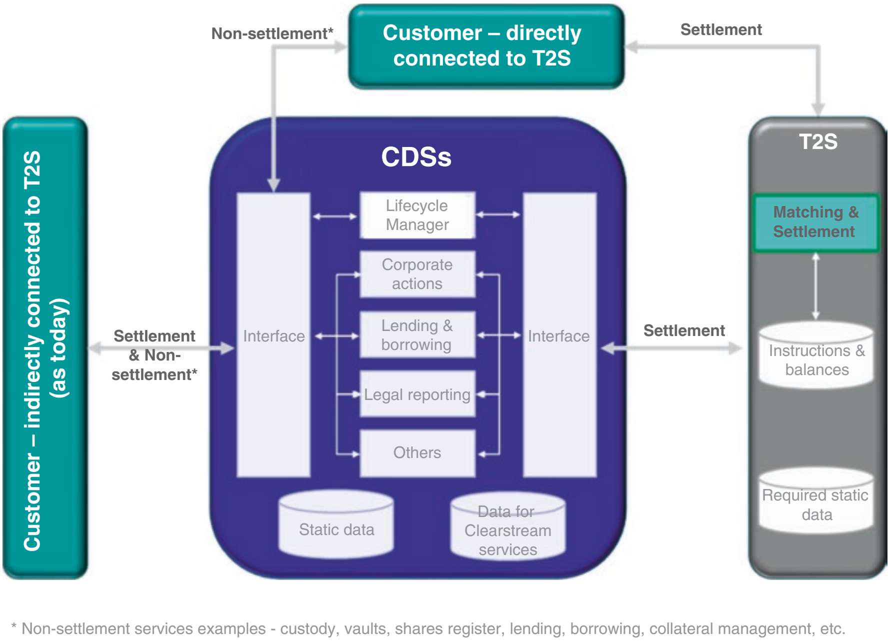


Fig. 17.2 Overview of activities outsourced to T2S (Clearstream Banking Frankfurt, April 2011)

Due to the fact that T2S is accommodating both accounts, securities and cash, the T2S parties (i.e. all who are holding accounts in T2S) can use a single cash account (the "dedicated cash account", DCA), and link it to several securities accounts ("SAC"). As a result, all cash proceeds out of all securities transactions conducted in T2S in CeBM as well as all income proceeds will be netted on a single CeBM account on T2S, the DCA. In turn, each DCA has to be linked to a real-time gross-settlement (RTGS) account held with a national central bank (NCB), since funding of the DCA occurs solely via the linked RTGS account. Hence, DCAs are "owned" by the NCBs, while the SACs are "owned" by the CSDs.

A comparison of today's situation with the one at the time once T2S is implemented shows the following main differences:

- Today, CSDs act as a quasi-monopoly. All domestic settlement and new issuance are centralised with the CSD. T2S will force CSDs into competition. Settlement services are no longer a differentiator. Issuers are no longer bound to use their home market CSDs. As a consequence competition will increase and the differentiation will focus on value-added services.
- Today, each market has its own settlement day schedule with individual (market) deadlines. T2S implements one harmonised settlement day schedule with a common deadline for all market players.
- Today, national barriers only allow few cross-border settlements in CeBM; settlement between customers of different CSDs is mainly processed on a free-of-payment

basis. T2S, in turn, negates this difference: settlement of a domestic transaction will be exactly the same as a settlement for a cross-border transaction.
- Today, CSDs operate their own settlement infrastructure, while T2S is the single-settlement infrastructure to which all participating CSDs outsource their settlement activities.
- Today, market participants maintain proprietary links to CSDs using Swift ISO15022 or CSD-defined standards in their communication. With T2S most CSDs will maintain this access path in order to shield their community for major adaptation costs. In addition, market participants will have the opportunity to connect and instruct directly to T2S. This option requires participants to use the new ISO20022 messaging standards, developed due to T2S and supposed to become the new standard for the securities business, and to maintain a dedicated network with a network service provider.
- CeBM is currently limited to the home central bank services. Connection between the CSD and the NCB is not standardised and cash and securities are not booked at the same time in the same system. In T2S accounts for both cash and securities are held on the same platform, allowing immediate settlement finality (real-time transfer of ownership).

Furthermore, T2S can be considered as an enabler for further harmonisation. Besides areas closely related to settlement such as messaging standards or the settlement schedule where we see harmonisation, there is also pressure on areas which are not directly impacted by T2S. Asset service remains under the full control of the CSDs. However, there are related transactions in the asset service area which will be processed in the T2S system. In order to allow a smooth cross-border settlement common rules on corporate action procedures have been agreed. As a consequence, markets are now adjusting to the standard. For example, the German market is now going to apply the record date, which is the date as of which holdings are entitled to take part in the event. Expectation is that this will lead to a significant reduction in market claims with harmonised processes. The ECB is monitoring the compliance to the agreed standards as one of the access criteria CSDs will have to meet in order to be allowed to migrate their settlement process to T2S. There are also asset services resulting in cash payment (e.g. income payments, redemptions) which will be booked in T2S. Going forward, all securities-related and -initiated cash proceeds shall be booked in CeBM to the DCA in T2S, rather than the current praxis to credit these to the RTGS accounts with the respective NCBs.

Another act of harmonisation which seems to be unrelated to T2S is the reduction of the settlement period (i.e. the time between trade execution and the settlement of the trade) to 2 days (“T+2”). In late 2014, almost all market moved from T+3 to T+2, thus reducing the exposure by 1 day. Those who have not done so are scheduled to do this prior to their migration onto T2S. One benefit lies within the treatment of corporate actions. Rights issues and deadlines can be harmonised as all settlements and entitlements are following the same settlement period.

CSDs will enter into competition in areas where they currently enjoy a monopolistic advantage: settlement, asset services and issuance of new securities. T2S enables CSDs to settle foreign securities in the same way as they currently settle their domestic securities—at the same ECB T2S tariff! T2S allows CSDs to act as a single point of access to all T2S markets. However, as services for securities settlement are harmonised, CSDs can no longer easily distinguish their services from each other. Focus will be put on collateral services, asset services and new issuance.

For the first time, there will also be competition amongst the NCBs for CeBM services. DCAs have to be opened via the NCBs, whereby a DCA can be opened with any NCB; it does not have to be the one where the related RTGS account is held. We will see situations where a DCA opened with one NCB will be linked to the RTGS account opened with another NCB. The function of RTGS account will be "reduced" to a funding account, since all relevant cash transactions in T2S will be booked in the respective DCA. As a result, service offering from different NCBs may vary.

T2S also has an impact on all CSD customer segments, i.e. investors, issuers and agent business. For them, there are more options to access the T2S markets. CSD customers, who currently are bound to "their" home CSD, can:

- Continue using an agent bank provider. Then these market players will access T2S markets indirectly and settle in commercial bank money (CoBM), with the respective exposure to the agent bank.
- Concentrate all assets with one CSD, who will then access the other markets and CSDs to enable settlement in CeBM.
- Become a "directly connected participant" (a DCP) at T2S, which allows them to send their instructions directly to T2S while the contractual relationship to a CSD remains unaffected.
- Become an "indirectly connected participant" (an ICP) at T2S, with more than one CSD. Here they would use the respective CSD for a given market (Fig. 17.3).

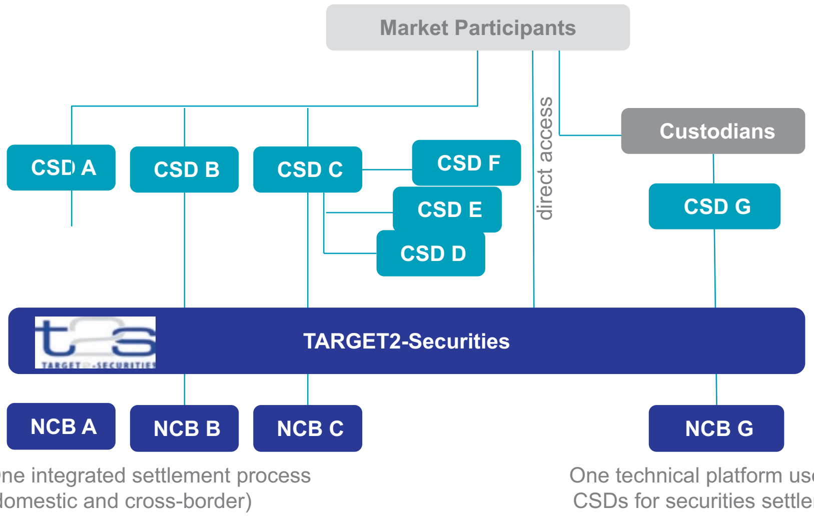


Fig. 17.3 Access possibilities for market participants in T2S

Issuers are currently issuing their securities in the markets where they expect most of the investors to be. As a result, a (e.g.) Spanish issuer addressing mainly Spanish investors is issuing his or her securities in Spain; an international issuer addressing a wide range of investors is more likely issuing his or her securities in international markets, such as the ICSDs. T2S enables issuer to reach his or her target investors through a single window, which is then selected due to superior services rather than proximity to the issuer.

All transactions settling in T2S are subject to the same fee schedule. The ECB is charging the CSDs at a publicly available price schedule, which shall stay unchanged up until at least 2018. There will be the same fee for both domestic and cross-border settlements. The aim of the ECB fee schedule was to set the price at such a rate which allows payback of the investments made over a certain period of time. The CSDs, however, are free in their fee strategy. It remains to be seen how different prices for settlement are going to be once T2S is fully operational, since some CSDs have already announced to only pass on the T2S charges to their customers and not to charge a premium.

This clearly shows that one pillar of the CSDs' revenue generation will disappear over time: the settlement or transaction fee. T2S puts pressure on CSDs as some may not be able to improve the services which remain with the CSDs in the same way as the ones for settlements are harmonised. Cost pressure will increase further and may lead to co-operations or further outsourcing of services between the CSDs. It is also very likely that CSDs may look for new revenue opportunities which will come due to the possibility to offer a single access to all T2S markets or even to move up the value chain and offer services which are today value-added services offered primarily by agent banks (e.g. tax services, proxy voting). With T2S, there will be choice for those entities wishing to access T2S via a CSD; the national monopolies will cease to exist. Although this is obvious for investors, the very same applies to issuers.

The aforementioned analysis conducted by Oliver Wyman revealed that T2S could allow brokers, asset managers and banks to take full advantage of the T2S model and realize savings between EUR 30 and EUR 70 million annually if they take action now to optimize their securities and cash supply chain. The study produces quantitative case studies showing that market players could realize significant capital, funding and operating cost savings by delayering settlement-related exposures, pooling collateral for settlement and tri-party purposes, netting more cash settlements and simplifying their operations (Fig. 17.4).

This and similar research activities show that there is lot to gain, provided that the industry players stop considering T2S as an IT project,[6] but rather an opportunity to streamline and adapt their T2S operating systems and create new products and services.

T2S will dramatically change the post-trading landscape in Europe. Customers can decide to appoint one CSD to cover all T2S markets, and thus can concentrate their securities with one single provider. This automatically leads to a reduction in

[6] Global Custodian/Deutsche Bank Survey, GC Magazine Fall 2014, p. 66.

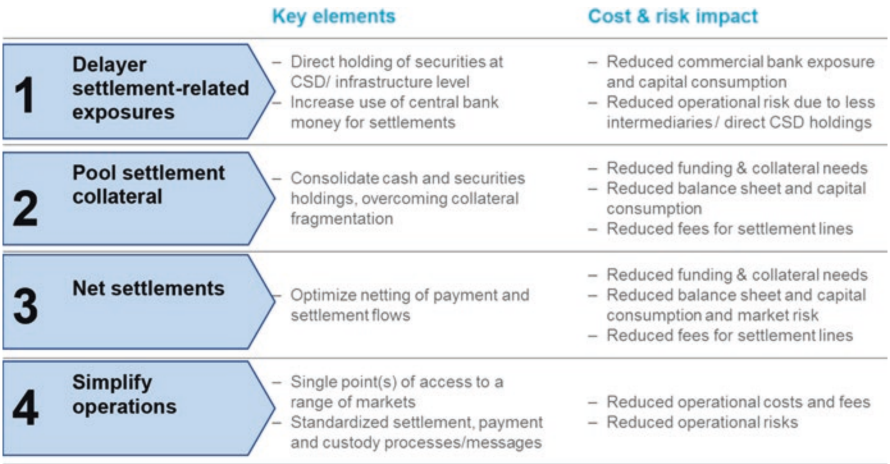


Fig. 17.4 The four efficiency levers for market participants to optimize post-trade economics (Oliver Wyman, September 2014)

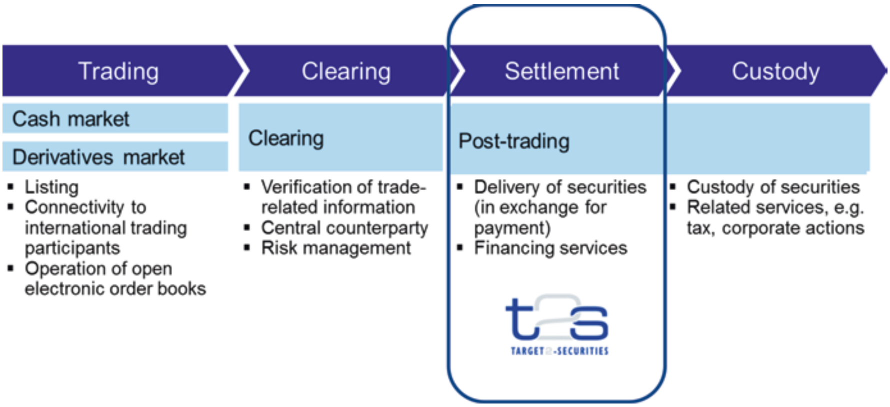


Fig. 17.5 From trading to custody (*source*: Deutsche Börse Group)

the collateral needs. T2S parties can link a single CeBM cash account (DCA) to various securities accounts (SAC), thus achieving a pooling of CeBM liquidity. This should even further reduce the need for collateral.

Besides the changes T2S is stipulating on the post-trading area, it also has impacts on the whole value chain (trading, clearing, settlement). The creation of a settlement platform with harmonised services actually leads to opportunities for trading and clearing. The number of multi-listings is likely to increase. Trading of T2S eligible securities is becoming more attractive, since settlement will be harmonised. Trades executed on a stock exchange in one country can then settle in a CSD of another country. Central counterparties (CCPs) are no longer restricted to have their clearing members holding accounts with the same CSD: T2S is laying the grounds to process cross-CSD settlement instructions initiated by the CCPs for or on behalf of their clearing members, e

ven when they hold the accounts in different CSDs. Last, not least, CSDs can place themselves as settlement location for trades executed on exchanges which are linked to T2S for settlement. Customers can work with their CCP of choice and use the account with their CSD of choice to settle all T2S-eligible securities in CeBM. Organisations which are covering the whole settlement chain are in the position to offer their customers all services out of one hand (Fig. 17.5).

Chapter 18
IT in Transition

Philipe Enness and Andrew Graham

Financial markets are currently going through a substantial period of transition, probably the most significant in living memory. This is leading to a fundamental review of the historical business models of all participants from exchanges, sell- and buy-side firms along with vendors (both data aggregators and application providers), suppliers of base technology and value-added services. No participant is left unaffected. To refer to it as a transition could be a major understatement, revolution maybe a more appropriate metaphor. These changes are ongoing and, when the transformational journey is complete, the markets will without doubt look very different than they do today.

Information technology (IT) itself is also going through an evolution with the emergence, adoption and importance of cloud, analytics, mobile, social data, distributed shared ledgers and cognitive computing. It is only just being understood in financial markets and other industries. Organisations are rapidly adapting their business models to reflect these changes from delivery, provision of services and pricing. The combination of these events is potentially overwhelming, but it is also rewarding for those organisations that take advantage, adapt, deliver and innovate around the challenges both from a business perspective and from an IT perspective.

The pace of change is accelerating, new disruptive start-ups are emerging and they have access to investment and low-cost technology that enables them to deliver faster than was ever possible before, new capabilities at an ever-lower cost.

Both author equally contributors for this book.

P. Enness
Global Lead Markets Infrastructure, IBM Banking and Financial Markets,
IBM Singapore Pte Ltd., 9 Changi Business Park Central I, Singapore 486048, Singapore
e-mail: philipe@sg.ibm.com

A. Graham (✉)
IBM Distinguished Engineer, IBM Banking and Financial Markets,
IBM Singapore Pte Ltd., 9 Changi Business Park Central I, Singapore 486048, Singapore
e-mail: andrew.graham@uk.ibm.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_18

Traditional organisations that fail to adapt may find themselves on the margins at best, and most probably irrelevant as we move further into the twenty-first century.

18.1 Three Critical Factors

Three critical factors in particular underlie the changes that have transformed, and will continue to transform, the industry.

Firstly, the degradation of the banks' balance sheets as a result of the financial crisis of 2007/2008: This has had a significant impact on the cost of capital and the return on equity.

Secondly, the tsunami of policies, directives and regulatory interventions in response to the financial crisis and in its aftermath—the erosion of customer confidence, the desire to capture and manage systemic risk and the political desire to correct perceived mistakes of the past: The estimated impact on the industry by

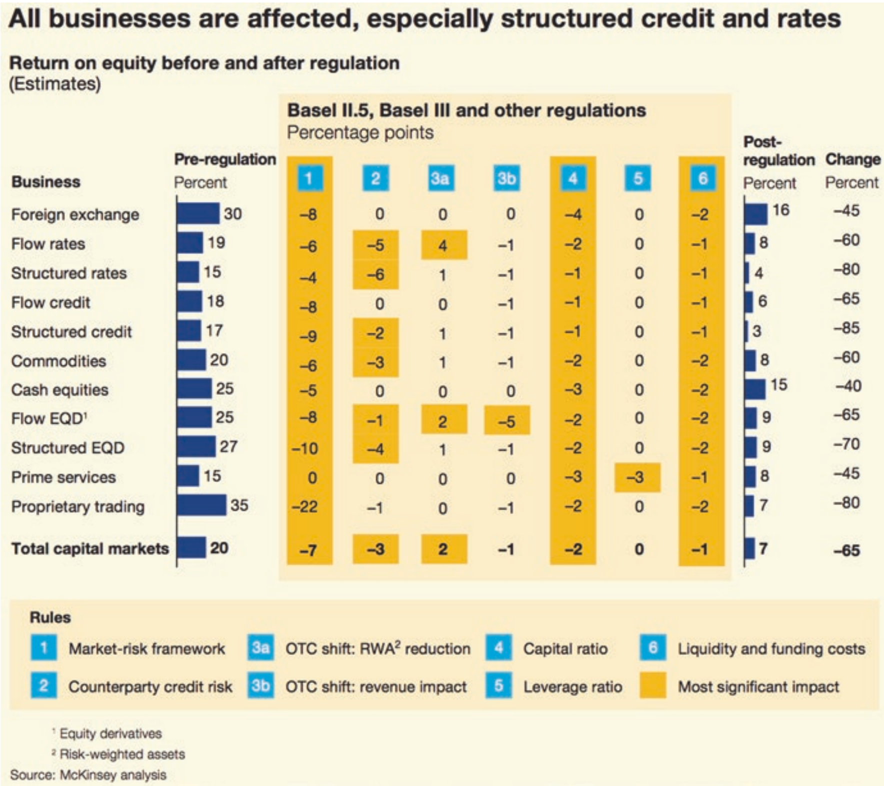


Fig. 18.1 All businesses are affected, especially structured credit and rates (McKinsey & Company - Global Corporate and Investment Banking: An Agenda for Change, 2011)

asset class (before any remediation efforts) can be seen in Fig. 18.1 taken from the McKinsey & Company - Global Corporate and Investment Banking: An Agenda for Change, 2011. It highlights the impact these events are having on the industry and why participants are having to review their business models and seek new revenue and aggressively manage costs.

Thirdly, the critical challenges surrounding the pace of change in technology (covered later in this chapter). In particular, the way technology itself is delivered, the explosion of big data and the use of mobile devices are impacting both financial markets and other industries as well. Overlay the technical challenges that are emerging with the impact of existing and more traditional technology developments of today (primarily around the reduction of latency, field programmable gate arrays (FPGAs), graphic processing units (GPUs), network cards, switches, etc.), coupled with some of the more recent business developments, market fragmentation (multiple execution venues co-listing same or related symbols). These innovations have both enabled and empowered high-frequency trading (HFT). They have also encouraged the use of direct market access, combined with low-latency connectivity, co-location and proximity hosting which in itself has led to an argument around "good and bad liquidity" and the fairness of markets. As witnessed by the heated debates following the release of the book "Flash Boys" by Michael Lewis, the future can only be more complex and disruptive.

18.1.1 Implications for Securities Exchanges

Exchanges are not immune to these challenges. Indeed, in their capacity as central intermediaries, they are viewed as a mirror on the broader financial market industry. Consequently, they are now positioned as vehicles to manage and reduce systemic risk and, as the wave of new regulation comes the way of the financial markets, exchanges in particular have become a major area of focus. Some examples include SEC Rule 623 (Consolidate Audit Trail or CAT) in the USA, Markets in Financial Instruments Directive II (MiFID II) and the International Organisation of Securities Commissions (IOSCO) principles/guidance that includes trading fee models, their impact on trading models, margin requirements for centrally cleared derivatives, financial benchmarks and exchange-traded funds (ETFs).

Exchanges have also become vehicles for the industry to implement elements of regulatory rules. This includes over-the-counter (OTC) derivatives moving on exchange, the requirement for exchanges to assist their clients meet their own regulatory challenges via additional reporting and a more timely delivery of information.

Exchanges have had to adopt technology in order to survive and innovate in the marketplace. Many have spoken of being technology organisations first, and exchange organisations second. Exchanges have invested heavily in technology and in acquiring technology companies to deliver innovation and value to their customers, as well as to diversify their business mix. This has not come without a price however. Significant costs have been added to the cost base, and the challenge going forward will be the ability to reinvest and continue to innovate when the margins of previous years are no longer available. Combining this with an ever-shortening life cycle, for both products and services that are deployed, it becomes clear that exchanges have to reinvent the way they deliver technology and services in order to survive in this competitive marketplace.

18.1.2 Concerns on Systemic Failures

The pace of change and the rapid deployment of new technologies do not come without risks. As exchanges negotiate this period of technology change and transition, they are increasingly having to respond and adapt to the identified challenges that could potentially impact their ability to protect against new and unidentified threats—more specifically, to systemic risks which are growing in complexity and are more difficult to anticipate. Indeed issues continue to surface, as the following illustrates.

18.1.2.1 Flash Crash of May 6, 2010

On May 6, 2010, the prices of many US-based equity products experienced an extraordinarily rapid decline and recovery. That afternoon, major equity indices in both the futures and securities markets, each already down over 4 % from their prior-day close, suddenly plummeted a further 5–6 % in a matter of minutes before rebounding almost as quickly. Many of the almost 8000 individual equity securities and ETFs traded that day suffered similar price declines and reversals within a short period of time. Shares were falling 5 %, 10 %, or even 15 % before recovering most, if not all, of their losses. Some equities experienced even more severe price moves, both up and down. Over 20,000 trades across more than 300 securities were executed at prices more than 60 % away from the values they had just moments before. Moreover, many of these trades were executed at prices of a penny or less, or as high as $100,000, before the prices of those securities returned to their pre-crash levels. By the end of the day, major futures and equities indices recovered to close at losses of about 3 % from the prior day.

18.1.2.2 Nasdaq's Facebook IPO

While all eyes were on the market on May 19, 2012, the day of Facebook's IPO, the exchange that handled the listing, Nasdaq OMX, was facing major challenges. Nasdaq knew that this would be the largest IPO and that individual participation would be unprecedented. Its systems to handle this type of trading had been tested in preparation for this event. However, a major unexpected "glitch" caused havoc for hours. Following the Facebook IPO, Tabb Group's "IPO Survey: Market Barometer" found that the impact of Facebook's IPO on investor confidence was almost as great as the Flash Crash.

18.1.2.3 Knight Capital Group

On August 1, 2012, a "minor change" to Knight Capital's trading software caused a major crisis for the firm. The technology in its market-making unit affected the routing of shares of about 150 stocks to the New York Stock Exchange. The chaos lasted less than an hour, but in the initial 8 min Knight accumulated an $8 billion position. The cost to Knight Capital of approximately $440 million in pre-tax losses nearly destroyed the company. The firm was subsequently acquired by Getco Securities.

18.1.2.4 Everbright Securities

On August 16, 2013, Everbright Securities, a Chinese broker-dealer, initiated a surge in trading volume on the Shanghai Stock Exchange. The index rose 53 % and reached its highest levels since 2009. Everbright Securities had placed accidental buy orders for billions of Yuan worth of ETFs, resulting in inflated orders totalling 23.4 billion Yuan. China's Securities Regulation Commission (CSRC) fined Everbright Securities 523.29 million Yuan ($86 million) and banned the company from proprietary trading in stocks and derivatives after the trading error disrupted markets.

18.1.3 A Myriad of Challenges

Once layered over each other, it's easier to see the myriad of complexities the industry faces. At stake is not only the industry's ability to prosper, but for the organisations and structures of today to survive in any resemblance of the marketplace we know. If we further overlay the emergence of cyber security as a significant threat, we can see that exchanges have major challenges and opportunities from the shifting business landscape, the specific technology demands and the wider disruptive technical challenges that are transforming all industries.

Given these challenges, new financial market business models are being developed by all industry participants. The result is a surge of players moving into territory that was previously the exclusive domain of the sell side or buy side and/or exchange/market data vendors/aggregators. Even technology providers are trying to differentiate themselves, forming alliances and partnerships that are blurring the old demarcation lines (Fig. 18.2).

The marketplace is slowly coming to terms with lower volumes, reduced fees and smaller margins. The "New Normal" is leading to fundamental shifts in participants' target operating models and their responses to the required changes are still being analysed and executed. Participants in their search for revenue growth are

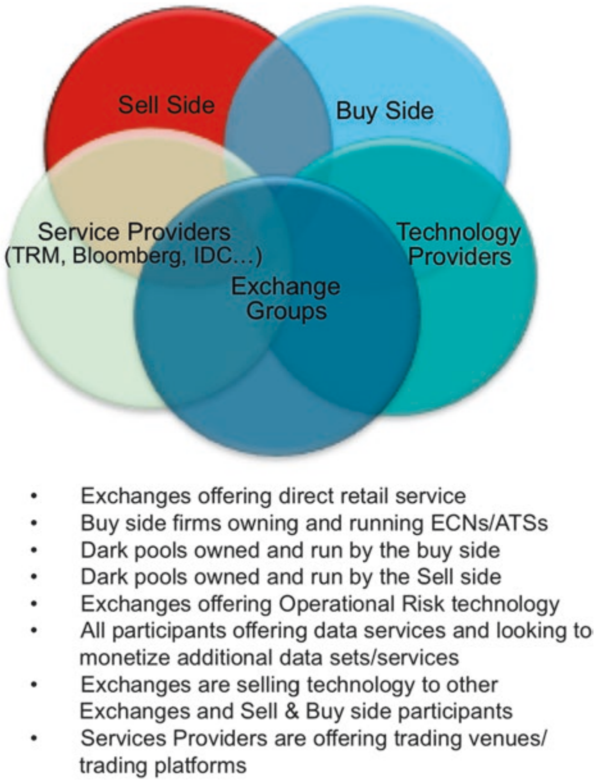


Fig. 18.2 The demarcation lines are shifting

looking at all areas of activity, both traditional and non-traditional. This shift will see coopetition grow, and new alliances will be formed. Technology is a significant driver of this change, as it is allowing organisations to transform their businesses at a faster rate with a lower investment cost than is normally associated with offering new services. Building alliances and an ecosystem will be a key component for achieving success as coopetition brings new challenges to participants, and those that achieve this difficult balancing act will be the successful firms of tomorrow. Exchanges find themselves in a difficult position as they are at the centre of the current ecosystem and have the most to lose if they get the balance wrong between new revenue and alienating their existing customer base.

18.2 Technical Drivers

18.2.1 What Are the Technical Drivers Shaping the Future of the Industry?

It is an extraordinary time to be working in industries that have technology at their core. The rate of innovation is growing exponentially and when multiple synergistic domains converge we experience an even steeper pace of change. Exponentials are hard for most people to picture; nevertheless, in a matter of years, industries can change radically due to the impact of rapid progressive innovation or even faster from disruptive innovation. As I write this, the UK Government is set to approve driverless cars on the roads in early 2017—a great example of many disciplines of technology coming together to create a synergistic value that has far-reaching social, ethical, technological and commercial consequences. And this will probably happen much sooner than anyone predicted only a few years ago. Technology trends have traditionally focussed on faster, cheaper, smaller aspects. Nowadays, the focus is firmly on agility and competitive advantage. I highlight a number of significant technology-fuelled changes that are on the horizon. Taken together, they accelerate the exponential impact:

- The threat of decentralised business models driven by peer to peer and distributed shared ledger technologies
- Empowering users and allowing everyone to be a developer
- Automation of technology—the complexity grows but is hidden
- Machines grow in power through AI, Machine Learning and Cognitive capabilities
- Risks increase from a security and complexity viewpoint
- Standards drive innovation
- Data as the new oil

18.2.2 The Threat of Decentralised Business Models

The Internet, a collection of open standards for communication, is driving a democratisation of everything that can be digitised, from music and films to finance. Technology-driven networks are replacing bureaucracy-driven hierarchies. Take, for example, how peer-to-peer business models for lending and equity crowd funding are disrupting traditional businesses. Emerging crowd models for loan syndication are appearing, further attempting to disintermediate existing value chains. The relatively recent invention of the Bitcoin & Blockchain-derived platforms—global distributed decentralised cryptography-based asset registers is a major threat to the status quo at country, government and global levels. For the first time in history technology makes it possible to transfer property rights (such as shares, certificates, digital money) at speed, in a transparent and very secure manner without the need for a central trust body. This technology has potential for far-reaching impacts to custody and exchange business models. This new technology has the promise of providing a mechanism for programmable money too—a concept where the business rules governing the value of exchange are contained within the digital system itself. We will see the appearance of new asset classes, new business models and new peer-to-peer services. One such example is the growth of the LMAX Exchange for FX, a break-off from Betfair that disrupted the retail gambling market.

18.2.3 Power to the User Where Everyone Is a Developer

This is the age of self-service, and users want to be empowered to execute change and to realise insight without relying on others, especially IT departments. Also, the user community has increasingly grown up with technology and, along with the consumerisation trend, they expect things to just work. The new researchers, for example quantitative analyst, analyst, trader and economist, are also "developers". They are familiar with spreadsheets and technology such as R, Python and Matlab. Coding will be blended with a world where users compose applications using rich tools. Users will drag and drop data, analytics and services together, then point the output to a browser or an SMS gateway or trigger a message to submit an order. See Fig. 18.3 for an early DIY Composable Analytics system called Beacon from IBM. Technology will be much easier to use, and it will empower users to act as developers without the complexity that exists under the covers.

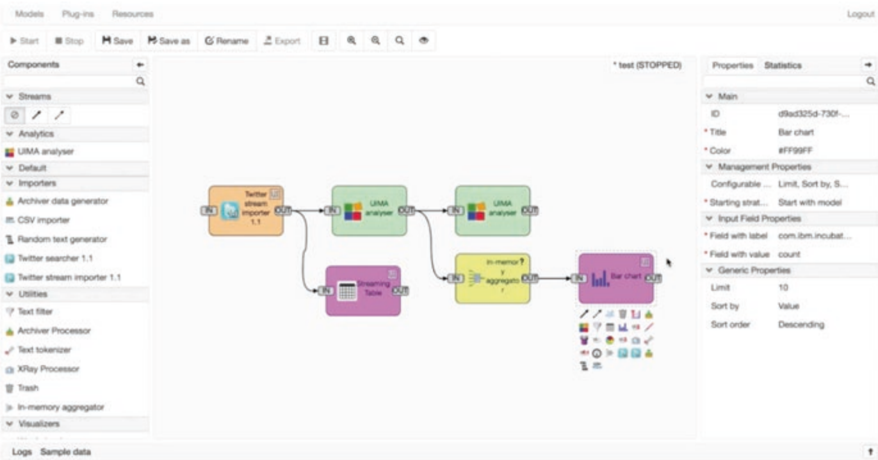


Fig. 18.3 DIY composable business applications

18.2.4 Automation of Technology: The Complexity Grows but Is Hidden

The underlying complexity of technology will continue to grow—the number of servers and the millions of lines of code to get things done will increase as will the volume and complexity of data around the world. However, consumerisation and the empowerment of users require automation. Examples include automation to provision IT, automation to recover and automation to suggest the next best action or question to address. This automation is driving cloud computing and the developer operations (DevOps) revolution which is a response to the interdependence of software development and IT operations. Technology is delivering on the promise of agility at a scale and underlying complexity never seen before. Users will expect a server to be provisioned in minutes, a database available shortly afterwards, and data to be populated a few moments after that. There is no room to wait for someone to install a new server, arrange power or arrange networking.

18.2.5 Power to the Machines

Machine learning, cognitive computing, artificial intelligence and human-computer interfaces will be a disruptive force in the coming years. Technology will be able to assist and guide users, and may be able to replace certain user tasks over time. This also has far-reaching social consequences as more jobs get automated and robotised.

18.2.6 Risks Increase from a Security and Complexity Viewpoint

This underlying complexity potentially increases risk. On the one hand, human decision making is reduced, and on the other hand the sheer number of underlying parts will be greater. Resilience will occur, for example, in the applications and in scale-out in-memory data stores rather than in the infrastructure. The threat of cyber security (and security in general) will be an enormous risk going forward, especially as the rewards for criminals increase. The promise of very secure crypto-assets may have a place to play here, as well as the ability to perform business logic on encrypted data without the need for decryption. Geopolitical risks are likely to increase, and thus technology will have to plan and speedily adapt in an increasingly unpredictable world. It will have to do this from an operations viewpoint, and from a data viewpoint, ensuring compliance with global and local regulatory demands.

18.2.7 Standards Drive Innovation

Standards (especially open standards) drive innovation. Examples such as Linux, HyperLedger, OpenStack, ISO 20022, FIX, BIAN and FIBO are examples of such standards—a form of crowd sourcing. Standards can be used for control reasons, but when open and useful can foster innovation and allow complementing firms/users to innovate. This is especially true right now with public APIs exposing a firm's high-value services.

18.2.8 Data as the New Oil

Bringing these various trends together, it usually comes back to data: good data, appropriate analytics, new insight and trusted outcomes and actions. This is true for social businesses, and the insight they gain from contextual data.

Data from the "Internet of things"—the emergence of billions of new connected Internet devices—can act, for instance, as alternative financial data, and as proxies for identity and credit worthiness. Data impacts application architectures going forward, and the role of the data scientist—who will be empowered to create new business value.

18.3 Enterprise and Retail Behavioural Convergence

There is a fundamental shift in how enterprise is buying and deploying technology across all industries, and the financial markets are in no way immune. It is a shift not driven by IT departments and their CIOs, but by their own customers and business users.

Two key drivers are fuelling this fundamental change.

The first is an outcome of the perennial debate between IT departments and their business users regarding "value for money". Desperate to remove significant cost from their operations, businesses are now seeking to take ownership of the acquisition of IT services directly. It is a seismic shift that is directly impacting the IT department's reliance on traditional revenue and margin models.

The second is considerably more fundamental, and the consequences are only just being understood. Driven by the success of the Apple iPhone™, Apple iPad™ and Apple App Store (a concept that has been embraced and exploited within the retail space), enterprise buyers are now expecting to replicate this experience and take advantage of the low cost of entry, and the ease of accessing complex services via the cloud. It is a concept that has rapidly transformed a number of industries; the heavily regulated financial markets, however, have been slow to respond to this trend that is here to stay, and that will impact all areas of business going forward.

Indeed, this trend was reinforced by a recent global survey by consultancy firm Avanade [2] titled "B2B is the new B2C: Work redesigned: Seizing New Opportunities in Changing Workplace". One of the key insights documented was that 60 % of companies that have built new business processes and technologies to accommodate shifts in customer interactions have reported increased revenue. Sixty percent also report a larger customer base, and 61 % report a growth in customer loyalty.

With the power shifting firmly to the new generation of users with their different attitude to technology and risk, it is the first movers and the innovators who are reaping the benefits of this behavioural shift. One needs to only witness those companies that are translating traditional smartphone games into business-oriented means. By attracting users to their platforms, they encourage them to become more familiar with complex products. Ultimately, this drives real transactions and economic flow.

It is a new generation of users who will continue to expect an experience that they are familiar with and who will demand more from their providers. This will impact financial markets along with many other industries. Those organisations that can deliver these experiences in the context of their industry will be the successful firms of tomorrow. The emergence of the Bloomberg App Portal allowing access to their historically proprietary data and environments is a good example of a traditional business reflecting the new model.

18.4 Technology-Fuelled Business Models

18.4.1 What Are Technology-Fuelled Business Models?

Alternative approaches are needed to support the "new normal" in terms of the anticipated rate of change that businesses are expected to achieve. Called for are new means to harness data, to offer new analytics and to deliver new insight. A new approach is needed where the innovation does not just come from within the organisation.

Open APIs, the API economy, complementors, standardisation, new cloud operating environments and business platform thinking are shaping the business landscape bringing innovation, agility and new opportunities. New entrants are embracing these forces, and faster than ever they are able to compete against traditional business models.

Inspired by studies done by Annabelle Gawer [3] from Imperial College in business platform thinking, there is a very useful model of how businesses can become more agile and more innovative and can adapt faster than their competition. The key

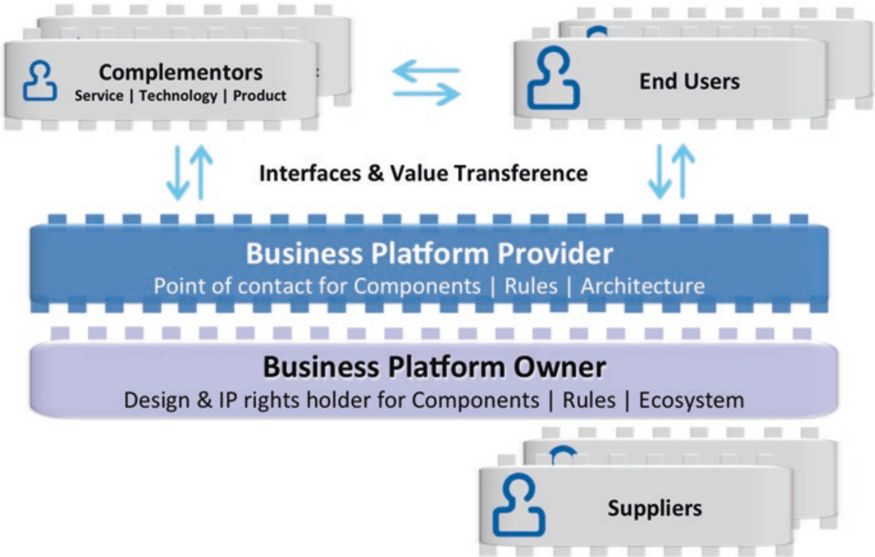


Fig. 18.4 Disruptive business platforms [4] and the importance of complementors

to their success is opening up high-value services through open interfaces/APIs and embracing "complementors" (Fig. 18.4).

Adam Brandenburger [5] of Harvard and Barry Nalebuff of Yale identified the role of complementors in their 1997 book called "Co-opetition". They highlighted a gap in Michael Porter's 1979 framework familiarly known as Porter's Five Forces. Porters Five Forces included the threat of the entry of new competitors, the intensity of competitive rivalry, the threat of substitute products or services, the bargaining power of customers (buyers) and the bargaining power of suppliers. Nowadays, based on Brandenburger & Nalebuff insight, complementors are often referred to as the Sixth Force.

18.4.2 What Are the Key Functions Within a Business Platform?

The "Platform Core", comprising owner and provider functions, should bring value to the overall system. It should perform at least one essential function, enable innovation and novel use, be easy to connect to through open interfaces, be easy to build upon, enable variety, enable low-cost development, allow fast adaptation and evolvability, be stable, generally have low variety, have high reusability, typically offer economies of scale and be difficult to replace. In summary they open a core capability to be accessed by partners and customers.

"End users" are demand creators. They benefit from the ecosystem of complementors and the platform core. In a multisided market, end users are often looking to be matched with other end users—supply and demand matching.

"Complementors":

- Offer complementary services, products and technologies, accessed through known modular interfaces such as APIs.
- Benefit from the features of the platform core, the wider ecosystem, other complementors and ultimately the end users.
- Can add value and capture value themselves.
- Will more likely invest in a platform if there are many end users, and if the platform is sufficiently open.
- Must not feel threatened by the platform competing with them.
- Increase the value to end users in a number of ways, including stand-alone functionality and performance, the size of the install base and the availability of complementary goods.

The platform owner and provider usually contracts for a known product, service or technology with a supplier, and generally the supplier, although innovative, does not innovate in a way that a complementor would.

"Interfaces" provide all actors with a means to access a platform. The interfaces may be industry standard based (e.g. FIX, FpML, ISO20022, FIBO), but certainly need to be standard. The level of openness will determine how easy it is to access the platform, and which rules, regulations and contracts are in place to use the platform. The more open the platform, the more innovation there generally is from complementors. These are the APIs that are fuelling the new API economy.

18.4.3 What Are the Benefits?

By opening core high-value services as APIs, development kits, freemium data, commercial analytics or free visualisations as example techniques, end users and importantly complementors can innovate on top of the core platform.

The first major benefit is the network effect. Simply put, the more complementors there are, the more end users will be attracted to the platform, and the more end users there are, the more complementors will be attracted to the platform.

Secondly, it encourages the platform owner and provider to focus on their core competencies, and allows others to innovate increasing the perceived value of the platform. In the past platform providers have tried to satisfy as many end users as possible by developing products, technologies and services in-house, but now end users demand unique features and the platform provider cannot generally afford to satisfy this need for mass customisation.

Thirdly, the control remains with the platform owner, but the innovation for the "long tail" of requirements comes from others.

18.4.4 Practical Examples of Applying This Approach

- Exposing market data, APIs, reference data, client and market insight data to complementors to innovate on your platform—co-creating value
- Allow innovation through standardisation of the internal and external interfaces and analytical models
- Exposing data, analytics and services through application stores and marketplaces
- Embedding data, technology and services in other firms operating platforms
- Allowing end users and complementors to instigate changes themselves, at their speed, through new tools such as a composable application tooling to make the core platform more easily consumable to less skilled users
- Expanding into new markets through complementors and partners
- Gaining new insights from social platforms, such as Estimize, StockTwits, Twitter and eToro that expose data across the life cycle

18.5 Implications for Enterprise IT in Financial Markets

The following sub-topics are examples of the area that enterprise IT must focus on to enable them to adapt to the ever-changing environment described earlier in the chapter.

18.5.1 Fostering the Appropriate Enterprise IT Culture

Enterprises typically adapt and evolve in a progressive manner, and successful disruptive innovation is rare. The threat of disruption is more likely to come from small "FinTech" firms targeting a specific niche. It is imperative to instill a "fast fail" culture within traditional enterprise IT. The term "fast fail" is a cultural and operational term to allow projects and ideas that do not meet the necessary requirements to stop quickly, therefore limiting unnecessary expense. The term "intrapreneur" is useful to describe the internal entrepreneurship that is required to drive this new rate of change. Collaborative working, both inside and outside the organisation, with appropriate reward and motivational approaches is key to generating and executing new ideas at speed and scale.

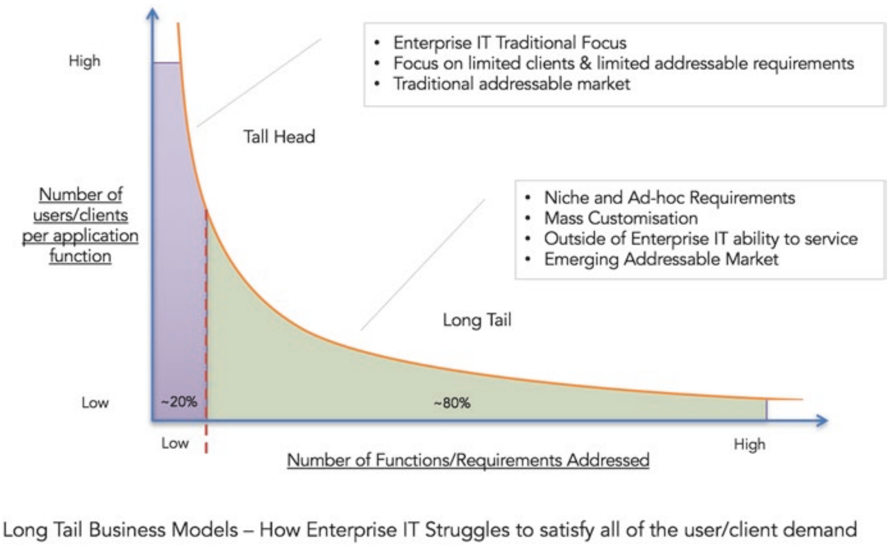


Fig. 18.5 Long tail business models

18.5.2 Progressive Openness Fosters Innovation

Enterprises typically have decades of locked-up embedded value in their people, processes, operations, analytics, services and data. As has been discussed, this needs to be opened up to new clients, new geographies, new partners and complementary services. To unlock this value in a controlled manner enterprises must open up their data, services, APIs, processes and value to their internal staff, their clients, suppliers, partners and complementors. Firms need to shift from product/technology-centric design to service/utility design focusing on user-centric thinking. For example, internal data should be accessible (in a controlled manner) to internal users through a service catalogue, as should service APIs. The new product manager should be the API product manager.

This opening up of data and services through standard interfaces, and standard data models, will foster innovation both internally and externally in the value chain. This in turn will drive competition and innovation, and will allow complementors to offer new services to the "long tail"—those end users and end-user requirements that internal enterprise IT doesn't have time nor budget to cater for (Fig. 18.5).

18.5.3 Everything as Service Through APIs

Internally, firms continue along their shared service strategies, and leaders are opening up their platforms to clients, suppliers and complementors. We are seeing shared services that include risk grids, know your customer (KYC), reference data, database, middleware, securities processing and regulatory reporting. In fact everything is considered as a service. Businesses are getting efficiencies through standardisation and sharing; over time we will see core business functions exploiting the new operating environment, and benefiting from the open API revolution.

18.5.4 User Empowerment, Automation and Agility

Users must be able to control their own destiny and not be overly constrained by enterprise IT's traditional rigidity. This will require significant amounts of automation in the enterprise, and pre-provisioned IT resources that predict users' demands for these services. Without automation, the expected levels of agility that users demand will not be possible. Automation also reduces cost over time and, done well, this offers a repeatable and controlled outcome that is vital in a highly regulated environment.

Extending these user services and data to mobile devices should be a checkbox service, a feature that is easily available and not a major IT project in itself. Users going forward will want to interact differently with IT, and will need progressive tooling to make their jobs more productive. For example, composable data and analytics tooling is currently an area of interest.

18.5.5 Framework and Platform for Data Movement

With a framework for openness and the underlying automation of IT in place, the focus must be on the data platform. In recent times, the "I" for information in the term "IT" has been suppressed compared to the technology element—but information is where the value lies. Data volumes and variety will increase and enterprise IT has to make data more consumable both within and outside the enterprise.

Currently data architectures are very Extract Transform Load (ETL) centric—meaning data moves around, is staged, passes through middleware, has value added to it and then moves again. The data architecture and platform for the enterprise will be a critical differentiator for successful firms. The latest thinking is grouped under the "data lake" terminology where data remains in place and is not moved around as frequently.

18.5.6 *Cognitive Computing and Machine Learning*

This new era of cognitive computing will have far-reaching consequences, and it is not easy to predict the outcomes. Whatever happens in the domain of cognitive computing, these systems will consume data and will publish data, and humans will, for a long time, remain the ultimate decision makers. So user-centric design will be critical. Focusing on the data platform now, and experimenting with new cognitive services, is probably a safe strategy for now.

Given the rate of current and expected business and technology change enterprises will need to look and behave more like agile innovative "FinTech" firms that partner with many complementary service providers to compete in this fast-changing environment. A focus on data, user empowerment, automation and ecosystem openness will be good foundational elements for ongoing differentiation.

References

1. McKinsey & Company - Global Corporate and Investment Banking: An Agenda for Change, 2011, Dominic Casserley
2. Avanade—'B2B is the new B2C: Work redesigned: Seizing New Opportunities in Changing Workplace'
3. Platforms, Markets & Innovation—Annabelle Gawer—Edward Elgar Press—2009
4. Imperial College—Platforms for integrated solutions—Miraglia and Davies 2009
5. Co-opetition—Adam M. Brandenburger and Barry J. Nalebuff—1996—ISBN 978-0385479509

Chapter 19
The Future of Finance: FinTech, Tech Disruption, and Orchestrating Innovation

Oliver Bussmann

19.1 Introduction: Finance and Information Technologies

At its most basic, money is not much more than a piece of information. Financial markets, in turn, are at heart sophisticated means of exchanging the particular kind of information money stands for. It is hardly surprising then that the financial industry has always had a keen interest in information technologies, going back to the days of the telegraph and ticker tape.

Over the last half century or so, the predominating information technologies in financial markets have been digital, and they have completely transformed how these markets operate. Yet, while these transformations have been dramatic, there is reason to believe that far more dramatic transformation is on the horizon. This too will be driven by technological advances, but there is a twist. Unlike in the past, the catalyst for change is coming not from within the industry, but from without, driven by the so-called FinTechs.

In this chapter we survey the current FinTech landscape and the areas of the industry where these companies are already causing disruption. We then focus on one particular development—the blockchain—we think has the potential for the most profound disruption in the ways that financial markets function. As an example of how the industry is reacting, we conclude with a look at how one major financial institution, UBS, is working with peers and the FinTech community to help foster innovation.

O. Bussmann (✉)
Bussmann Advisory, Hofstrasse 52, 6300 Zug, Switzerland
e-mail: oliver@oliverbussmann.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_19

19.2 The Rise of FinTech

A predominant characteristic of technology in the financial industry has been its closed nature. Financial platforms are by and large designed, owned, and managed by the financial service firms who use them, or by specialty providers catering directly to them. The result has been large, complex, proprietary systems which often do not provide an easy interface to other platforms. Where there is shared infrastructure—SWIFT is the classic example—it has almost always been built, managed, and used by industry participants.

This closed technological world began to be challenged after the financial crisis when an increasing number of start-ups as well as some well-established tech names, all with little or no financial industry background, began developing products and services aimed at the financial industry value chain in ways and at a scale not previously seen.

These new entrants were quickly dubbed FinTechs, and their rise has been dramatic. In 2008 FinTech attracted some 1.4 billion US dollars in investment. By 2013 that had more than doubled, to four billion dollars. In 2014 the amount tripled to nearly 12 billion, and estimates are that this year investors will pour close to 20 billion into these companies[1] (Fig. 19.1).

We can point to a number of factors which have contributed to this sudden rise. By shaking clients' trust in the financial system, the financial crisis left many people more open to nonindustry alternatives than had been the case in the past. The crisis also forced banks and other industry players to focus inward, rebuilding balance sheets and regrouping. During this period, technology investment has also generally gone towards meeting postcrisis regulatory requirements. This has not been a period conducive to fostering technological innovation from within, leaving the door open for outsiders.

Technological advances have also contributed. The rise of the Internet, mobile technologies, and more recent advances like cloud computing and big data provide

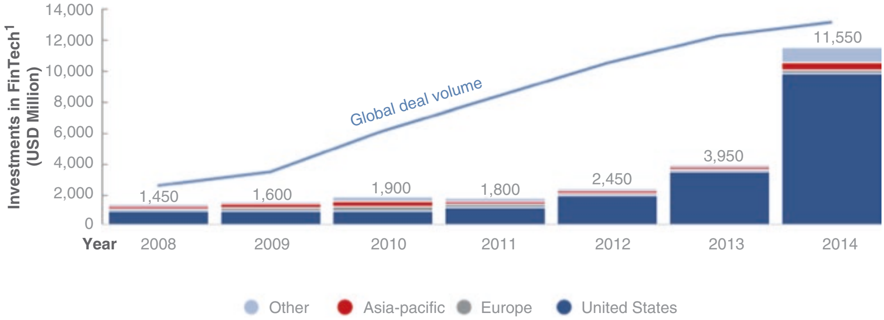


Fig. 19.1 The rise of FinTech: FinTech investment since 2008. Source: Accenture and CB Insight

[1] (Accenture and CB Insights. From UBS slide. Need to get exact reference.)

new opportunities for interacting with customers and carrying out financial transactions and related services. The general shift to more open systems, as well as the falling cost of technology, makes it easier to design and build platforms that can interact with each other and the existing infrastructure.[2] This is significantly lowering the barrier to entry.

Changing client behavior and expectations are a factor too. The more opportunity people have to carry out financial transactions through nonindustry providers, the more they get used to the idea. That in some cases these providers are well-known names they already trust, like Apple or Google, facilitates this process. As people transact more and more of their lives online, they increasingly expect seamless interactions with their service providers over a variety of devices and at all times. This is the kind of service that technology companies understand and have traditionally provided better than financial service companies.

Finally, FinTech firms have the advantage of agility. With no legacy systems to deal with, and, for the moment at least, largely unregulated, they are able to start from scratch. This freedom to maneuver fosters innovative thinking, as does the collaborative, experimental atmosphere typical of technology start-ups, where all ideas are welcome and failure is seen as a virtue, driving the process forward.

19.3 FinTech: Areas of Disruption

The potential disruption of the rise of FinTech has not been lost on incumbents. While it is not uncommon in the media and elsewhere to portray the rise of FinTech as an existential threat to banks and others, in truth many in the industry welcome these developments as a catalyst of innovation. FinTech brings added competition, but also the potential to provide tools and methods to help banks deal with the cost and regulatory and market pressures they have been facing in the postcrisis world.

In 2015, after more than 15 months of research, the World Economic Forum (WEF) published a major survey of FinTech disruption under the title of "The Future of Financial Services."[3] The report was developed in collaboration with a number of major industry players—including UBS, the organization for which I work—and over 100 innovative new entrants and subject market experts.

The result is the first consolidated taxonomy for disruptive innovation in the industry, as well as probably the most comprehensive road map of the current and potential impact of FinTech on financial services published to date. As someone who was personally involved in the process, I can attest to the fact that the results are rather eye-opening (Fig. 19.2).

[2] See among others "The Fintech 2.0 Paper: rebooting financial services," Santander, InnoVentures, Oliver Wyman, Anthemis group, 2015.

[3] The Future of Financial Services. How disruptive innovations are reshaping the way financial services are structured, provisioned and consumed. Final Report, June 2015. www.weforum.org.

Fig. 19.2 FinTech areas of disruption for financial services. Source: WEF

Take payments, for example, one of the most fundamental services that the industry offers. Technology firms have been moving into the payment space, offering consumers more varied and easier ways of making purchases than incumbents. While Apple Pay and Google Wallet are among the most well known, they are hardly the only technology companies to insert themselves between banks and the end customer in this way. This is driving the move to a cashless world, with interesting ramifications.

On the one hand, it is making the payment process transparent and automatic for consumers, potentially disrupting other payment channels, in particular credit cards. On the other, as more payments are carried out via mobile devices and the like, there are more opportunities to capture customer data of the kind that can be very interesting for merchants and others, including credit institutions. This is driving interest in these systems from the vendor side.

FinTech is also very active in the area of cross-border payments. Today it can still take hours and even days to send money between financial institutions in different countries through the standard infrastructure. FinTech has already spawned at least two alternatives that do the job better. On the one hand there are decentralized systems using cryptographic protocols to transfer money, often via cryptocurrencies like bitcoin. These generally rely on distributed, trusted ledgers that allow for safe, secure, fast, and inexpensive transfer of value directly between parties, bypassing the middleman. (We examine the distributed ledger in more detail below.) There are also mobile money schemes, like MPESA, that rely on a trusted central party but leverage mobile communications to, among other things, reach the unbanked in developing countries. Both approaches are creating competitive pressure on incumbent infrastructure.

Deposits and lending are another fundamental service provided by the financial industry that is now facing competition from FinTech. Good examples are the peer-to-peer lending platforms like Funding Circle or Lending Club, which aim to match savers and borrowers directly. They are not only bypassing banks, but also implementing completely new business models, for example using social media and other nontraditional sources to help assess credit risk and handle adjudication. With lean, automated processes, such platforms open up lending to a broader customer base. This creates new competition for the traditional savings and loan industry.

A similar thing is happening in the capital markets, with alternative platforms that allow organizations to raise capital through direct investment or by easily issuing equity or debt. Unlike in the lending space, these platforms are not likely at the moment to compete directly with incumbents when it comes to large issuers. The new platforms are however opening up capital markets to smaller, more risky entities, like start-ups, in novel ways. They are also employing innovative approaches. Seedrs, for instance, matches individual investors with seed-stage ventures. By only issuing equity if a certain threshold of interest is met, it employs a "wisdom-of-the-crowd" risk assessment strategy indicative of the social-based techniques made possible by both new technologies and mindsets.

Banks have traditionally provided a number of sophisticated, value-added services to help clients manage investments—everything from gathering, consolidating, and analyzing market data to developing and executing trading strategies to monitoring performance and risk. Today, a host of FinTech providers are looking to provide outsourcing services for different parts of this value chain. By leveraging cutting-edge technologies, they can often carry out these functions in ways that are better and cheaper than what incumbents can offer.

This has interesting implications. On the one hand it gives incumbents access to better and faster tools to help advise clients. On the other, it is in effect putting all the pieces in place to build an automated robo-advisor, cutting banks out of the investment management process altogether. Platforms like FutureAdvisor, for instance, are able to automatically implement an investor's asset allocation, taking into account his or her risk appetite, and automatically rebalancing as the market changes. Ayasdi uses big data to develop and test market hypotheses and develop trading strategies. Kensho automates the modeling of investment scenarios, and can provide real-time projection of performance based on different assumptions. These and other developments could potentially alter the role of the client advisor, shifting the emphasis towards relationship management, while leaving the machine to provide the advice.

FinTech is also invading the trading floor. This may not seem new: algorithmic trading dates back to the 1970s, and high-frequency trading to the 1990s. But these past innovations were generally about being faster. The new generation of machines is all about being smarter. Computer-based algorithms can now read the news, for instance, and pass on information about market-relevant "real-life" events to traders instantaneously. They can monitor and analyze social media to pick up on trends traditional media might miss. They can leverage big data to pluck useful insights out of vast pools of information, allowing traders to profit from a much broader picture of what is happening in the world.

With these tools in hand, human traders will certainly become more well informed. They may also become redundant, as advances in artificial intelligence and machine learning mean that trading algorithms can now learn from their experiences, and even their mistakes. In the future, the robo-advisor may be implementing its sophisticated investment strategies with the help of the robo-trader, all without human intervention and at a cheaper cost to clients.

While these are some of the areas of disruption most relevant to financial markets, they are by no means the only ones. As the WEF report makes clear, whatever the future of finance will be, it will certainly not be the same as finance's present.

19.4 Meet the Blockchain

Of the FinTech developments mentioned above, there is one in particular, the blockchain, which we believe deserves special attention. This is because we believe that the blockchain has potential to cause massive disruption, both in the limited context of how securities markets function and in a far-broader societal context.

The blockchain was invented in order to create bitcoin, the first viable, mass market peer-to-peer digital currency. The concept behind bitcoin was proposed in a succinct, nine-page paper published in 2008 by a person named Satoshi Nakamoto. (Nakamoto has remained anonymous to this day, prompting many to conclude that this is a pseudonym for a person or persons unknown.)

The goal of the project was to create a "peer-to-peer version of electronic cash [which] would allow online payments to be sent directly from one party to another without going through a financial institution."[4] From the outset, therefore, it was designed to bypass traditional financial service infrastructure (Fig. 19.3).

Bitcoin was not the first digital currency ever proposed, but it was the first to take off. Despite a number of ups and downs, it remains the most widely used digital currency today. As of this writing, some 14.5 million bitcoins are in circulation at a market capitalization of close to 20 billion US dollars.[5]

[4] See bitcoin.org/bitcoin.pdf.

[5] Source: Deep Shift. Technology Tipping Points and Societal Impact. WEF Global Agenda Council on the Future of Software and Society, Survey Report, September 2015.

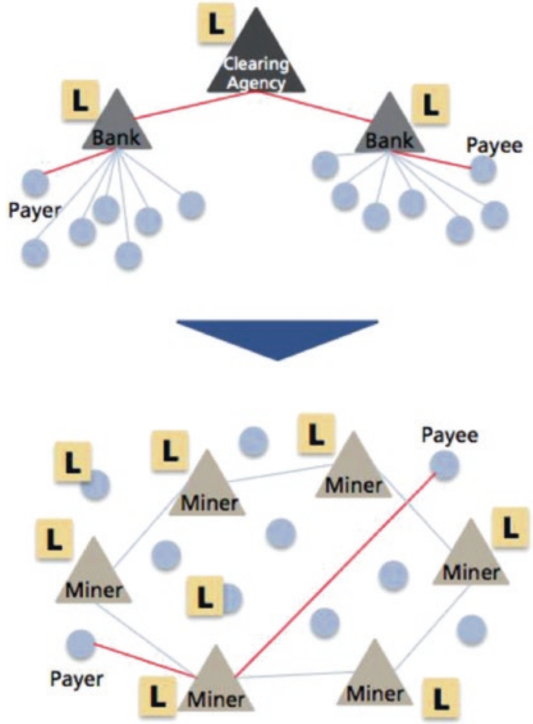


Fig. 19.3 As opposed to traditional models, which rely on a clearing agency, the blockchain uses a distributed, peer-to-peer network to transfer value. Source: UBS

For various reasons, including early adoption by criminal elements (drawn by bitcoin's ability to transfer funds anonymously and out of the reach of governments) and some well-publicized thefts, bitcoin has garnered a great deal of notoriety in its short life. This has tended to mask, at least in the public mind, what for many is bitcoin's most interesting attribute. This is not the currency itself, but rather the blockchain technology upon which it is based.

Irrespective of how it used, the blockchain has proven very successful as a peer-to-peer platform for transferring value (and much else) without intermediaries. That has serious ramifications, not least for those organizations whose business models are predicated on acting as such intermediaries.

The key problem the blockchain solves is not how to build a robust system for transferring value that does not make use of a third party. It is how to build one that can be used by people who do not know each other, and therefore, by definition, *cannot* trust each other. The answer it proposes is simple and elegant: remove the need for trust.

The blockchain does this by ingeniously emulating the basic functions of a trusted intermediary in an automated, self-verifying way. These functions include verifying the identities of the parties to a transaction, addressing and delivering transactions, creating a correct and immutable record of the transactions (a ledger),

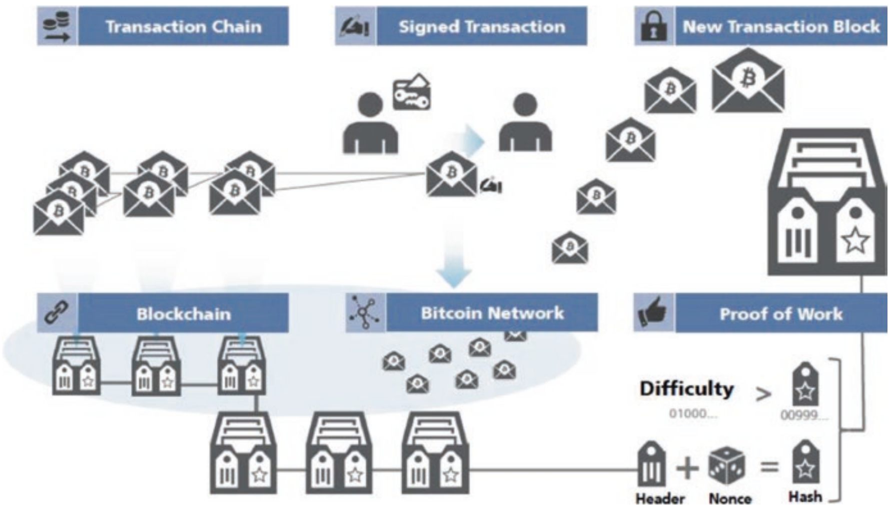


Fig. 19.4 A typical blockchain workflow. Source: UBS

and, most importantly, ensuring that cheating, through fraudulent transactions or double spending, is not possible.

While the concept is simple, its implementation is rather complex. It is perhaps best explained by looking at how a blockchain-based digital currency platform typically works (Fig. 19.4).

19.5 The Typical Blockchain Process

To begin, an interested party joins the platform and downloads an app. The app creates a digital signature uniquely identifying that party. The person then purchases or otherwise acquires the digital currency, perhaps by selling a product or service, which is credited to his or her account. (In reality there are no accounts in digital currencies, just entries in the ledger assigned to a particular identity, but the effect is the same.) When time comes to make a payment, the person creates a transaction in the app using the digital signature of the counterparty, which is also the counterparty's address.

This is where things begin to diverge from the approach taken by traditional systems. Instead of sending the payment instructions to a central transaction agent, the app simply broadcasts the transaction to the rest of the currency's peer-to-peer network (which anyone with the requisite technical know-how can join).

Using cryptographic techniques, all the miners on the network who receive the transaction then verify its authenticity. That is their first task. Their second task is to add the transaction to the ledger. To do this, each individual miner adds the transactions it has verified to a pool it keeps locally. Once a miner has a set of 500 verified transactions ready to add—referred to as a block—it broadcasts this block to the network, much as an individual would broadcast a transaction. The block is then received by all the other miners.

The next step is to add the verified block to the existing ledger, which is nothing other than a long chain of verified blocks going back to the first transaction—hence the name blockchain. Since the blocks in the chain are connected to each other using cryptographic hashes, the chain is extremely robust. No matter how long it is, if someone tries to alter even one character in it, its cryptographic properties will change and the fraud can be easily discovered.

The problem is that, since the blockchain does not reside on a single repository but instead is comprised of all the copies of the genuine chain in existence everywhere, it is imperative that there is agreement on what the correct chain is. That means agreement on the order in which blocks get added.

To achieve this, miners "vote" for which of the currently unassigned blocks gets added to the chain next. This voting process involves a complex and, importantly, computationally expensive "proof-of-work" process. Because this process costs CPU time and hence electricity, it effectively means that people have to pay to vote.

This is a fail-safe designed to make it prohibitively expensive and very risky for anyone to try and take over the system through multiple voting. Honest miners, on the other hand, are compensated for their efforts by being rewarded with new bitcoin for being the fastest to successfully process a block onto the chain, which provides the incentive to participate in the system.

Once a block has been added to the chain, all the other miners update their system with this new, verified version of the ledger. Because the "pay-to-vote" system, as well as other technological hurdles, makes cheating extremely difficult if not impossible, the end effect is the creation of a distributed record of transactions which cannot be tampered with, which updates itself without the aid of a third party, and which can be trusted even though none of the parties know each other.

Blockchains become even more interesting when, as is the case with more modern versions of the protocol, you can add instructions to the transaction along with value. This ability to specify conditions, for example "only pay A if B is received by C," allows for the creation of self-verifying "smart contracts." While the example above is rather simple, there are platforms being developed to allow for extremely complex business logic to be added to the chain. As these instruction sets, once they are in the chain, cannot be tampered with, they are guaranteed to carry out their instructions to the letter.

It follows then that, if you can find a way to connect real-world objects with entries in such contracts and ledgers, the blockchain becomes a robust means of automatically and cheaply transferring ownership of assets of almost any kind. This suggests a number of powerful new possibilities.

19.6 The Implications of the Blockchain

We believe that blockchain technology could be the biggest disruption in computing and finance since the Internet (Fig. 19.5). A public, open-source ledger increases transparency into transactions while providing mechanisms, for instance digital pseudo-identities, which can protect individual privacy. As many people have recognized, distributed ledgers and smart contracts can be adapted to authenticate and securely record a wide range of real-world financial transactions. They can also lower barriers to entry for new players, and will almost certainly drive industry innovation, to the benefit of clients.

What is by no means clear at the moment is exactly when and in what way this technology will be adopted. As in the early days of the Internet, much is in flux. However, the community is already considering an incredibly wide variety of use cases, both within the financial industry and without.

Among the more general uses, the blockchain could conceivably be employed to ensure secure digital identities, protecting consumer privacy and helping fight fraud. It can provide irrevocable proof of the ownership of assets, whether digital or, as ways are found to connect items in the ledger to real-world assets, physical things. This could help combat counterfeiting and copyright infringement.

Governments could profit from secure, network-based record keeping to expedite such processes as transferring title in a real estate transaction or carrying out secure online voting. Because the transaction ledger is fixed, the blockchain could make it much easier for governments and others to collect transaction-related infor-

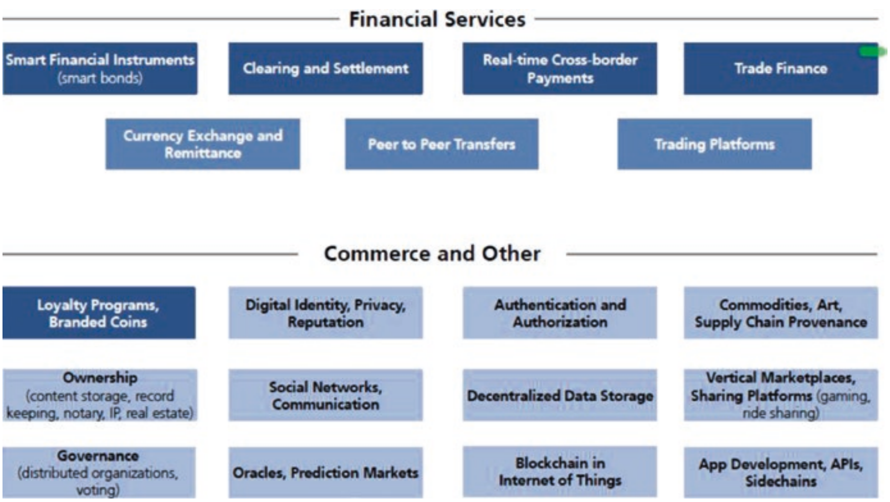


Fig. 19.5 While the blockchain is likely to disrupt financial services in several areas, its potential uses are much broader. Source: blog.ventureradar.com, letstalkpayments.com, CTO Research Services, UBS

mation. This could give regulators a real-time view of what is happening in markets, allowing them to implement real-time systemic risk management procedures. Economists could find themselves suddenly working with much more detailed, much more accurate economic data.

The blockchain will also be a great enabler of the Internet of things, providing a secure, trustworthy, auditable way for connected objects to communicate with each other autonomously. For businesses, it has the potential to greatly simplify processes, which will make for more efficiencies and more opportunities for low-cost automation.

This certainly applies to financial markets, where the blockchain has the potential to dramatically simplify processes associated with a wide range of transactions, such as clearing securities or transmitting international payments.

If we look at equity markets, for instance, we can see that in the primary market a blockchain-based system could automate equity issuance, creating securities with their own immutable digital identities. It can be used to automatically register ownership of a security, and could easily handle life cycle events, like dividend payments, stock splits, buybacks, or mergers. It can be used to record corporate actions, and to easily, and automatically, carry out and record transfer of ownership.

Things become even more interesting when we look at secondary markets. While in its current form the blockchain is still too slow to be used directly for trading equities, with its ability to provide direct and irrevocable records of transactions, it has the potential to radically simplify and speed up clearing and settlement processes. A securities market in which the current T+2 days' settlement regime becomes T+2 seconds is a very different market indeed.

By potentially eliminating many if not all manual processes associated with clearing and settlement, the blockchain will drive significant operational efficiencies. Since the ledger is decentralized and distributed all over the network, it is basically fail-safe, as a verified backup is always at hand. That means it can reduce operational risk. Among other things, this could simplify and reduce the cost of business continuity planning.

Direct, for all intents and purposes, instantly cleared transactions dramatically reduce settlement and counterparty risk as well. That might allow banks to take risk off their balance sheets, which could in turn lower capital requirements. Blockchain could significantly improve capital efficiency for other market players as well, as quick and efficient settlement, and the elimination of payment windows and cutoff times, means that invested capital is more quickly available for other uses. There will also be less need for posting (expensive) collateral.

Blockchain technologies can lower the cost of accessing the market, making it easier for smaller participants to deal directly. It also significantly lowers the cost of entry for newcomers looking to provide services to market participants.

The open nature of the distributed ledger can theoretically vastly simplify and improve compliance procedures as well. With digital identities, it will be easier to perform know-your-customer and anti-money laundering checks. If regulators have access to the chain, companies will be able to design regulation-aware processes, significantly reducing compliance costs and removing uncertainty.

Business administration could potentially be vastly simplified too, for example through the automatic payment of VAT. Areas like trade finance, which still rely on predominantly manual processes, could quickly become much more efficient. A smart contract could for instance easily execute a payment the moment a container, outfitted with an electronic sender, reaches port and is successfully opened by the recipient.

It is important to stress that, at the moment, these are all only possibilities. There are still a number of hurdles to be cleared before the technology gains wide adoption.

Blockchain proper has issues of speed which need to be overcome. There are also questions around scalability: a distributed ledger recording every transaction in a securities market would get very big very quickly, potentially overwhelming systems. There are external issues as well. The protocols around creating digital IDs and linking blockchain tokens to real-world assets will have to be worked out, and agreement on asset class workflows will be needed in this radically new environment. Similarly, the legal community and regulators will need to be highly engaged in the process, as we create a new fabric within regulated markets.

We believe that the key prerequisites for blockchain to reach the tipping point will be sufficient volume in platforms, robust security, and above all general trust in the system. We also strongly believe that this can be achieved only through cooperation between the industry, technology companies, and regulators. When all the stakeholders are on the same page, we think that the technology stands a very good chance of taking off and delivering on much if not all of its promise.

19.7 UBS and the Blockchain

For this reason, the industry is currently taking blockchain very seriously. UBS provides a case in point, and I think can serve as an informative example of how industry incumbents are evaluating and embracing the innovations that the blockchain potentially represents (Fig. 19.6).

Under a special program we began in 2015, dubbed "Crypto 2.0 Pathfinder," we have been both experimenting with the blockchain for our own uses and actively contributing to the development of open standards. Like many of our peers, we believe strongly that in this new world, collaboration with the community, as opposed to developing proprietary systems, will be crucial for success. For this reason, instead of carrying out these experiments in-house, we have been doing so at the famous Level39 incubator in London, which we joined in 2015 (the first global bank to do so). This gives us the ability among other things to interact directly with the community at large, share our insights, and profit from the insights of others.

Two early experiments which we carried out in our lab are indicative of how blockchain can be employed in real-world conditions.

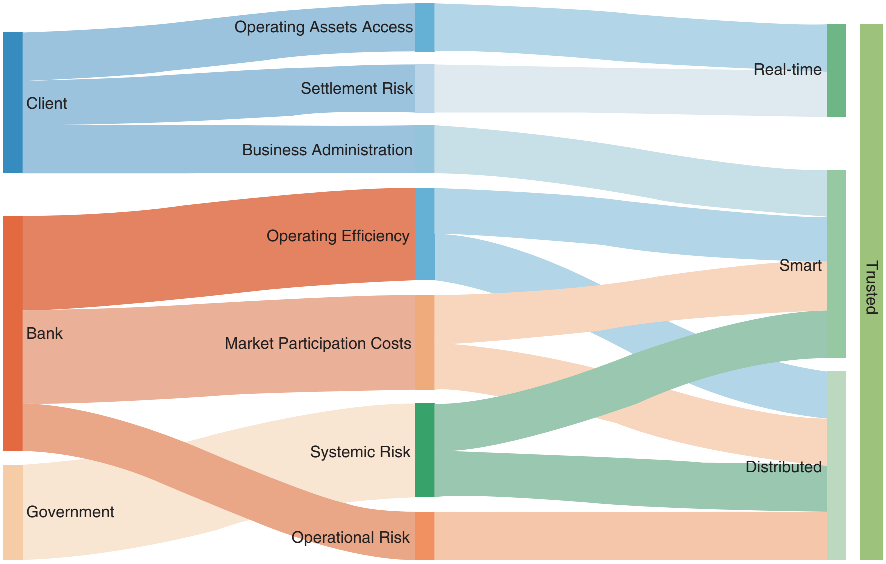


Fig. 19.6 Clients, banks, and the state could all benefit from blockchain technology. Source: UBS

In one experiment, we created a "smart bond" to validate the feasibility of the overall blockchain approach as well as our initial smart contracts hypothesis. To do this, we built an application on the Ethereum platform that can recreate a bond's issuance, interest calculation, coupon payments, and maturation processes. In our model, the issuer and the buyer do not need the support of pre- and post-trade intermediaries. The blockchain and smart bond contract handle the flow of information and money automatically and almost instantaneously.

Achieving this required the creation of our own virtual coin, which we called the BondCoin. Although dubbed a "coin," it is not actually a currency but a token, intended to be linked to real-world currencies and connected to a central bank account. Spending is therefore equivalent to spending real money.

We found that this experiment confirmed many of the use cases already mentioned. For our clients, such a platform would clearly mean a more convenient way to issue bonds, with lower administrative cost and increased speed. For the regulator, the platform would allow real-time visibility into securities positions and therefore a more precise view of systemic risk. In addition the regulator would be able to sign the business logic of a smart contract in advance, guaranteeing that the product would meet suitability requirements. For ourselves we found that a smart bond would mean reduced risk from real-time clearing and settlement as well as reduced clearing and settlement cost. If the regulator were to join the network the platform could also easily automate our regulatory reporting.

In a further development of the BondCoin, we conducted an experiment with a "utility settlement coin." Here the idea was to link the token to a real fiat currency on blockchain and use it to settle transactions carried out on financial platforms based on the technology. This promised a number of benefits similar to those seen in the smart bond experiment, but applicable to multiple asset classes. In a typical scenario, UBS might have its own blockchain-based platform to issue bonds, and another bank might have a blockchain-based stock trading platform, but both would use the same utility coin for settlement. This helps bridge the gap between asset classes and between the ledger and the real world.

19.8 Conclusion: Collaboration Is Key

Amid all this technological change, it can be easy to lose sight of the significant cultural transformations that are taking place in our industry as well. Along with new capabilities, FinTech is also bringing a new mentality to the industry. Banks and other industry participants are beginning to adopt the kind of new economy, collaborative mindset that has long characterized technology start-ups. They are also beginning to see great opportunities in working together, leading to more open, transparent, and decentralized approaches. This is transformative.

To cite one example, in September 2015 UBS joined a group of the world's leading banks, including Barclays, Commonwealth Bank of Australia, Credit Suisse, Goldman Sachs, JP Morgan, State Street, and RBS, in a partnership with R3, a financial innovation company based in New York, London, and San Francisco. A large number of other institutions have subsequently joined as well.

The intention behind this collaborative partnership is to design and deliver advanced distributed/shared ledger technologies to global financial markets. The idea is to develop commercial applications, but also to establish consistent standards and protocols for this emerging technology across the financial industry in order to facilitate broader adoption and gain a network effect.

To conclude: FinTech in general, and blockchain in particular, is still in its early days. However, the potential is certainly there to create an exciting, new, and radically more efficient and robust financial system. Technology is one part of the puzzle. Collaboration and openness are another. Together, they are creating the conditions for a major step forward for the financial system as a whole, and for all those who benefit from its products and services.

Chapter 20
Equity Capital Market Expectations of Corporate Issuers: The Fresenius Perspective

Ulf Schneider

20.1 Introduction

Fresenius is a global healthcare company with more than €20 billion in annual sales and more than 200,000 employees worldwide. During this past decade, our sales have increased almost threefold from €7.1 billion to €20.3 billion. Our market cap has increased from €2.4 billion to €20.1 billion in the same period. Our Deutsche Börse/Frankfurt Stock Exchange listing and XETRA trading experience are key components of Fresenius' growth story. Healthcare markets around the world are consolidating fast and while the industry particularly lends itself to debt financing, access to equity capital is a necessity to manage our overall financial risk. Over the last 10 years alone, we went through three capital increases from approved capital to finance growth projects.

20.2 The Importance of Efficient Equity Markets

A consolidated, liquid, and transparent equity capital market is not only beneficial in providing a reliable source of equity funding. Equity capital is also helpful as an acquisition currency and we have taken advantage of this as part of our growth story. Equity also serves as an important yardstick in the design of long-term incentive plans for our senior executives to align their interests with those of shareholders. A transparent equity market is important in this case as well—as in the cases of issuing equity or paying with equity as acquisition currency there needs to be faith in a fair and transparent price-building mechanism. Pricing ambiguity diminishes the use of equity for these purposes. Liquidity is a key requirement for winning large institutional

U. Schneider (✉)
Nestlé S.A., avenue Nestlé 55, 1800 Veve

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_20

investors as shareholders. These funds take sizeable positions in individual stocks. At the same time they want to be able to build and exit positions quickly without moving the market. Fragmentation of liquidity across platforms cannot be in the interest of listed companies, as it is not in the interest of their shareholders.

20.3 Towards a Global Shareholder Base

Over the past decade, the nature of equity investing in our company has continued to change. Gone are the days of trading being dominated by local long-only investors. German shareholders today represent less than 20 % of our free float. Our international investor base over the years has been steadily increasing. While the 1990s and 2000s saw primarily a larger number of investors from the UK and the USA, we are now seeing global ownership in our stock on the rise. This applies in particular to the large capital pools that went hand in hand and with the rise in the Far Eastern economies. As our company has grown and developed a global footprint, we also noted that interest in investing in our equity has grown with our presence in international healthcare markets. Market observers saw our success and found their way to invest in us through our XETRA listing.

20.4 The Trend Towards Professional Investing

Institutional investors today already represent 61 % of our shareholders. Daily trading volumes on XETRA in percent of the number of our shares have almost doubled over the past decade. The number of participants in our investor calls around the globe has quadrupled. The importance of derivatives for our share price development has increased significantly, in particular following our DAX30 inclusion in 2009. Today, there are more than 2500 certificates and about 3500 tradable options based on our company's stock. With them comes an improvement in liquidity, which facilitates getting into and out of our stock, thus providing market access.

20.5 The Relevance of Stock Indices

There has been a consolidation movement over the past decades when it comes to relevant share price indices. Industry-related indices are less important, whereas more focus is put on our DAX membership. This applies not only to derivative activity but also to benchmarking our performance. In particular, our joining the DAX in 2009 has left its mark. We have experienced more market capacity when issuing equity, more liquidity, more derivative trading activity, and even a reputation effect that transcended into the debt market and helped us finance our growth.

20.6 Efficient Markets as Public Goods

When it comes to the future of equity markets, corporate issuers are only indirectly involved. Short-term, they are passengers, not drivers. Nonetheless, corporate equity issuers have clear preferences and given a choice they will gravitate to markets that best meet their needs. Clearly, a fair and transparent price-building mechanism is at the top of their wish list. Information technology has revolutionized trading. Time delays and privileged access to information have been vastly reduced—notwithstanding the recent attention that high-frequency traders are getting. Fair and transparent price building has "public good" characteristics. While the fragmentation of trading and increased trading through so-called dark pools may offer marginal benefits to individual market participants, stock market regulators, corporate issuers, and the broader investing public should be careful not to risk losing the benefits of this public good. A case in point is to compare price building and trading in equity markets around the globe to the international trading in corporate debt. The latter has increased tremendously over the past few decades but is highly fragmented and nowhere near the transparency and trust in price-building mechanisms that the equity markets enjoy.

20.7 Building Trust in Capital Markets

In a day and age when the Basel III requirements could potentially restrict future corporate lending through banks, capital markets gain in relative importance as a source of corporate finance. A solid stock market presence with consolidated trading activity, transparent trading, and price building, combined with meaningful regulatory oversight, are the cornerstones of long-term capital market access. With the rise of securitized and traded debt, equity market success will be one contributing factor to debt market success and vice versa. As markets converge and information technology improves, debt and equity are getting more and more connected. Innovative stock market organizers with cutting-edge information technology and transparency, liquidity, and appropriate counterparty presence can help to make capital markets more resilient and reliable. As the various price-fixing scandals in other markets, such as Libor and foreign exchange rates, have shown, any loss in trust leads to downsides for all parties concerned.

20.8 Conclusion

Pooling capital flows, regulating them prudently, and making them transparent provide the public good mentioned earlier, i.e., a generally recognized, reliable, and trusted price that represents market participants' true view of value and is not subject to manipulation and partial optimization. This is the best service that regulators, exchange operators, and academic research around the globe can provide to us as corporate issuers to support our success in the long term.

Chapter 21
The Investment Process

Asoka Wöhrmann

21.1 The Investment Process

While being at the core of every asset manager's business, the investment process is a discipline that is still subject to continuous efforts for professionalisation and standardisation. The reasons are manifold. First, clients are becoming more demanding and the evaluation of the investment process is an integral part in their choice of the investment manager. Furthermore, in Germany an increasing number of institutional clients rely on consultants in their choice of an asset manager, a development pioneered in the Anglo-Saxon investment business. These consultants also put the quality of the investment process at the top of their priority lists, when it comes to evaluating asset managers. And finally, asset managers themselves recognise the benefits of a rigorous investment process for their business.

When at Deutsche Bank in 2012 various previously independent investment management entities were put together under the Deutsche Asset & Wealth Management brand, it gave the company the opportunity to align its existing investment process to the new structure. Establishing a unified active investment process supports one of Deutsche AWM's prime objectives: to deliver consistently strong performance to clients worldwide. It assures investors and their advisors that portfolio management is conducted in a consistent, controlled and transparent way across strategies.

There is a major challenge when establishing such a unified investment process for a company like Deutsche AWM. Especially in its equity business, Deutsche AWM follows a bottom-up approach, embedded in a holistic strategic house view. Investment decisions rely to a great extent on the know-how and

A. Wöhrmann (✉)
Deutsche Asset & Wealth Management Investment GmbH,
Frankfurt am Main 60612, Germany
e-mail: asoka.woehrmann@db.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_21

quality of the investment professionals. The investment process therefore has to balance the necessary leeway every portfolio manager needs to generate alpha with the aim to provide a framework that ensures consistency, control and transparency.

The four main pillars of Deutsche AWM's investment process are research and the CIO View, the portfolio construction, portfolio management and quality management. Each of the first three pillars is supposed to contribute to alpha generation.

1. Research and CIO View

   (a) Deutsche AWM's global research platform generates investable recommendations based on fundamental analysis. These recommendations are rigorously challenged and combined to generate a coherent global market view, the CIO View.
   (b) Research is undertaken by research analysts, often a dual function of portfolio managers.
   (c) The CIO View generation is driven by the CIO, supported by the CIO Group which comprises investment, research and asset class heads and is based on the expertise of our investment professionals.
   (d) Top-down-driven strategies can be expected to be in line with the CIO View.

2. Portfolio Construction

   (a) Portfolio construction is performed for each core strategy. Portfolio construction teams (PCT) include lead portfolio managers, who are held accountable for all investment decisions in lead portfolio.
   (b) The expertise and idea pool of the PCTs—which are designed to foster knowledge sharing and encourage a culture of open debate and challenge—contribute to superior investment decisions.
   (c) Portfolio objectives and guidelines are aligned with clients or distribution partners. The clear definition of product objectives and constraints creates accountability and focus. The relevant PCT knows about needs to be achieved, and clients know what to expect. The construction outcome can be measured and evaluated against the objectives.

3. Portfolio Management

   (a) Portfolio management is carried out by individually responsible and accountable portfolio managers. They take into account specific needs and objectives for each individual portfolio and bear the ultimate investment decision. Trade execution is done by dedicated specialists for the relevant market segments.

4. Quality Management

   (a) The quality management function is responsible for the continuous evaluation of performance, risk and skill or—more generally—strengths and deficiencies across the process and within process components.

(b) Quality management is a holistic approach aimed at measuring, evaluating and improving process quality in order to achieve consistent, repeatable results. A dedicated team ensures that evaluations are undertaken objectively and "at arm's length" from investment teams.
(c) Investment process management is part of quality management and examines whether the established process works as desired and whether the interfaces between the process components operate effectively. For example, IPM analyzes how and to what extent research and the CIO View are used in portfolio construction as well as the quality of recommendations from research and the CIO View.

In other words, the aim of the systematic, transparent investment process is to align the alpha-generating functions of the CIO View/research, portfolio construction and portfolio management.

The organisational structure does not necessarily reflect these various functions (e.g. investment professionals may perform both a research analyst and portfolio manager role).

An important characteristic of the investment process is its non-hierarchical approach. While the CIO View establishes a macro-framework the ultimate investment decision lies with the responsible portfolio manager.

The investment process is a dynamic, two-way communication process with an inherent feedback loop between the bottom-up and the top-down approaches.

21.2 Investment Decision

Deutsche AWM's investment approach is research driven and bottom-up oriented. Theme identification and, for the equity business, stock picking are the central pillars, as it is believed that inefficient capital markets offer opportunities for active managers to outperform.

The global research platform generates independent investment ideas through fundamental analysis and leveraging insights across asset classes. Proprietary fundamental research is the foundation for Deutsche AWM's investment decisions and underpins the entire investment process. Often portfolio managers hold dual roles and also serve as research analyst, generally as members of a global team for a certain sector (equity) or an asset class (fixed income). Worldwide, there are 150+ fixed-income and 100+ equity investment professionals.

The research platform comprises two major components:

(1) Macroeconomics:

The economists provide a thorough analysis of the world economy and the main economic regions. Through the macroeconomic research, major trends are identified and their impact on various components such as GDP growth, inflation and trade flows is analyzed.

(2) Fundamental analysis and value assessment:

Research analysts provide research and recommendations on equity (regions, sectors, size [small/mid caps] and single companies) and fixed-income markets (issuer, rates, currencies, duration, etc.). Investment rationales and drivers are discussed and a proposal for short-term (3-month horizon) and long-term (1-year horizon) investment assessments of the underlying market and security prices are produced.

As a signatory to the United Nations Principles for Responsible Investments (UN PRI), ESG is considered in the fundamental analysis.

These two components are the main inputs for the CIO View (strategic and tactical). The CIO View offers a directional world view that enables investment orientation internally and externally. The CIO View also has two components:

(1) The strategic CIO View provides a long-term outlook and orientation.
(2) The tactical CIO View represents a short-term assessment and positioning. It establishes a generic, investable framework of unconstrained recommendations. The tactical CIO View includes:
    (a) The global CIO recommendations (GCR) which are a subset of the overall global market view
    (b) A relative value asset allocation which provides relative value orientation between sectors and regions
    (c) Tactical vies are expressed in relative terms using a five-point scale from −2 to +2 (including 0) which is believed to allow for sufficient differentiation to convey the expected degree of direction and conviction. The short-term recommendation rating is accompanied by quantitative entry, target and review levels as well as a qualitative rationale.

Team discussions, knowledge sharing and challenging of investment theses are integral parts of the investment process, which is believed to foster good decision-making and lead to a more consistent approach across client portfolios within the same strategy. At the same time, the portfolio manager has the final decision on portfolio positioning and is responsible for portfolio performance (Figs. 21.1 and 21.2).

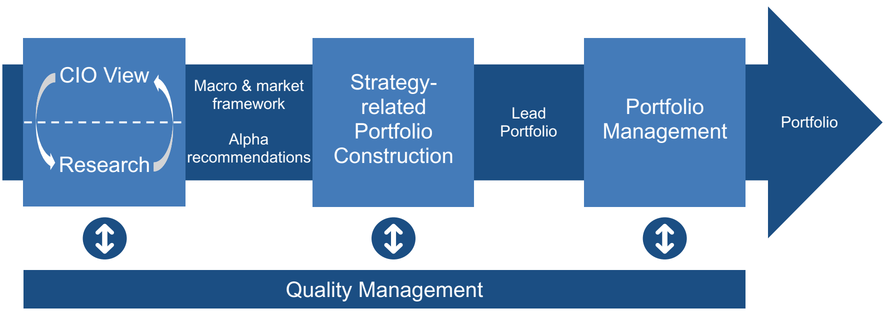


Fig. 21.1 Pillars of Deutsche AWM's investment process

Investment recommendations	Global Top Pick	IPO assessment
— Fundamental analysis summarised in stock and financial note — Recommendation (strong outperform, outperform, neutral, underperform, strong underperform) — Target price	— "Strong Outperform" classifies to become a Global top pick — "Fresh Money" idea — Minimum upside potential of 15%, minimum market cap. of EUR 1.5 billion	— New company assessment — Recommendation "Subscribe" versus "Not Subscribe"

G-Cube

Meeting	Frequency	Participants	Output
Equities morning meeting	Daily	All Equity PMs, analyst and chief economist/strategist	Discuss global macro and company news-flow
Lead Portfolio team meeting	Weekly & ad hoc	Lead Portfolio team members	Discuss portfolio and new investment ideas
ESG working group	Bi-weekly	ESG investment specialists	ESG methodology, integration into platform, etc.
Sector teams	Monthly	Global sector team members	Sector views and investment ideas
Global Top Picks list	Monthly	All equity PMs and analyst	Discuss Global Top Pick recommendations
CIO Equity View	Monthly	Global CIO, equity business heads	Macro handout, equity sectors and regional slides, tactical CIO View
CIO Day	Quarterly	Global CIO, chief strategist, Equity and FI business heads, regional and sector team heads	Strategic CIO View across all Asset Classes (Equity, Multi-Asset, FI)
Update on global regions and sectors	Semi-annual	PMs and analyst	Investment themes and stock recommendations

Fig. 21.2 Components of the CIO View

In general, portfolio managers are encouraged to take positions that lead to outperformance through personal objective settings and through the incentive compensation structures. The goal is to create an environment where portfolio managers have the necessary support and environment to employ their skills to achieve or exceed clients' risk/return objectives. The degree to which investment professionals achieve these objectives is being measured and they are rewarded accordingly.

The investment preferences of a portfolio construction team in a bottom-up-driven investment strategy, such as equities or high-yield fixed income, may deviate from the CIO View. If so, the differences are discussed and challenged in the PCT, which may lead to either the confirmation of a distinct bottom-up-derived positioning or an adjustment based on the appreciation of the drivers underlying the CIO View.

On a single portfolio level it is the portfolio manager who is responsible for portfolio implementation and ongoing management. Portfolio managers use the lead portfolios as orientation to manage the portfolios within a core strategy, thus ensuring that our best ideas benefit all clients.

PCTs and portfolio management teams exercise a stringent buy and sell discipline. Consequently, changes in the market environment or investment teams' convictions are rapidly reflected in portfolios. Examples of triggers that could cause a shift in portfolio composition are given in Fig. 21.3.

21.2.1 Asset Liability Management

As the majority of the funds from institutional investors that Deutsche AWM manages ultimately have to serve pension commitments, liability-driven investing (LDI) plays an important role within the company. This is because the long-term horizon of pension commitments coupled with the necessity to have specific amounts of liquidity available at specific times makes it mandatory to put the liability structure at the centre of the investment strategy.

What is LDI all about? LDI shifts the investor's mindset away from a pure return orientation. Instead, liabilities are used as a benchmark. The investment's performance, as well as its risks, is always measured against the value and structure of the liabilities. Generally speaking, LDI tries to smoothen the asset performance in order not to deviate too much from the liabilities. In that respect the goal of an LDI solution provider will be to help pension scheme trustees to better understand and address the risks they face in meeting their fiduciary responsibilities. LDI is not only about matching future cash in- and outflows. It addresses a wide range of other risks that may affect a pension scheme: interest rate risks, inflation risks and credit spread risks (Fig. 21.4).

There are various misconceptions about LDI strategies. We name three of them and explain their shortcomings:

Decision	Potential trigger
Sell	— Investment ideas identified — Significant increase in conviction ideas — Increase in risk allocation — Significant upside due to increased target price
Sell	— Fundamentals deterioration — Violation of risk budget — Identification of better opportunities — Price target reached
Adjustments	— Intact investment thesis but change of degree in conviction — Geopolitical topics might force risk and / or position changes

Fig. 21.3 Triggers that might cause a shift in portfolio composition

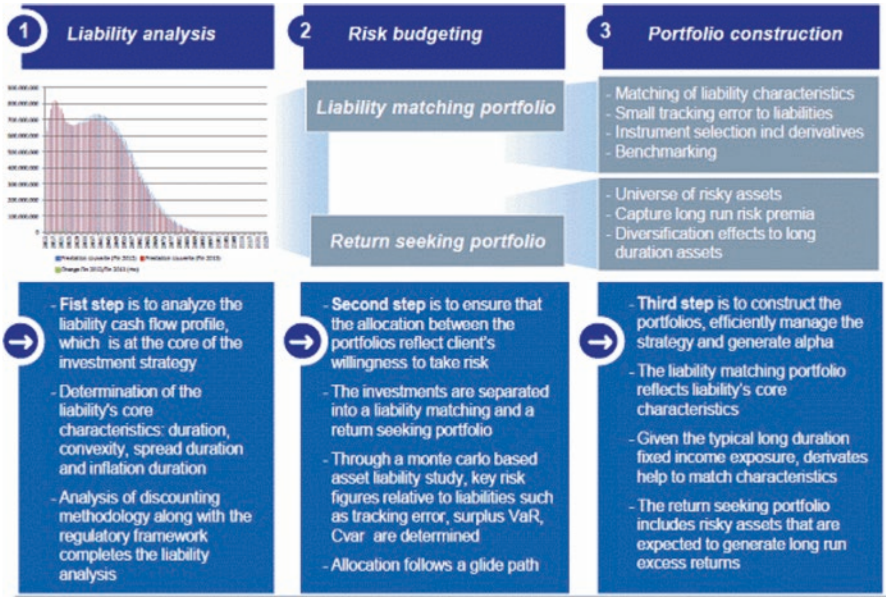


Fig. 21.4 Investment approach within the LDI strategy

- LDI simply means duration extension. Duration gaps against liabilities are often identified as a major source of risk, and the reward for duration risk is often low. However, it is just one of the factors driving the risk analysis. The analyses and quantification of all risks will reveal various ways to optimise the relationship of assets against liabilities.
- LDI avoids risk taking and market opportunities. LDI serves to take the risks that have the highest expected payoff, reduces unrewarded risks and ensures that risks taken remain within the defined risk budget.
- LDI should only be implemented when the pension plan is close to full funding. The LDI strategy includes the framework how the fund's strategy changes over time, taking the degree of funding and associated levels of risk taking into account. This supports the governance process and provides clear objectives for all stakeholders.

From a product point of view, managing third parties' pension schemes requires four different building blocks, all of which Deutsche AWM can provide either on a stand-alone basis or as an integrated solution: the trust, in Germany the Kapitalverwaltungsgesellschaft (KVG), the investment strategy and the asset liability management study. Deutsche AWM employs an entire dedicated team for LDI solutions. The full value chain of services that the client can expect comprises the risk budgeting, actuarial expertise, LDI oversight and reporting and derivatives management.

The ALM study stands at the beginning of the whole process. It is a labour-intensive exercise for which the supplier has to be able to draw on various models, data and legal frameworks that can take up to a year. It involves a detailed analysis of the pension scheme's asset and liability profile, return expectations, investment horizon, risk tolerance, variances and correlations as well as economic and legal restrictions. Based upon these findings client constraints can be established, the investment universe can be defined and a portfolio optimisation is conducted. Furthermore, various sensitivity analyses (in respect to inflation, interest rate for example) are carried out.

Elements of the ALM toolbox that Deutsche AWM provides to its customers are:

- Full capital market modelling (interest rates, spreads, inflation, credit migration and default)
- Implicit consideration of roll-down and mark-to-market changes of fixed income indices and liabilities
- Implementation of nearly any asset class (expected return/volatility) possible
- Implementation of market value-based management rules
- Flexible reporting of results (any market value-based risk figure could be generated)
- Tracking error, expected funded status, CVaR (95 %), VaR (95 %) and probability of additional contributions.

Once the framework is established, the investment approach within the LDI strategy involves three steps:

While fixed-income assets clearly still dominate LDI investment strategies, some of the pension schemes managed by Deutsche AWM have an equity ratio of up to 40 %. Considering the current low or even negative yielding bond environment, the portfolio mix of pension schemes is expected to undergo further changes.

Another dynamic development within LDI touches the two major forms of pension schemes offered by companies. Defined benefit plans (DBP) used to be the typical way in which corporates rewarded their employees. With the full risk of capital market developments resting with the companies, it made LDI a necessary risk-minimizing tool. While DBPs still dominate the pension world (and will continue to do so for some time, as even for closed funds the remaining time frame can cover a couple of decades), defined contribution plans are the dominant option for new pension schemes. But while companies benefit from the lowered uncertainties by only having to set aside a certain amount of money every month, it puts the risk burden on the employees' shoulders. In order to increase the predictability for the employees about their futures pension payments, LDI strategies are being used again.

Chapter 22
Institutional Investors and Exchange Organizations

Asoka Wöhrmann

22.1 Introduction

To begin this chapter, it is best to first take a step back and have a brief look on the last few years:

Financial markets typically follow their very own momentum in changes they undergo. Nevertheless, especially the years since the financial crisis took off in 2007 have been coined by an even higher speed in significant and longer lasting overhauls.

Sufficient discussions have been conducted and countless experts have been requested to share their view on the turmoil that financial markets went through in the subsequent years, but the public discussion only scratched some other fairly important evolvements.

Now, from my view, that is the view of a globally oriented and acting asset management and thus investment company, these developments require a closer examination in order to properly prepare for the future in financial markets. In this chapter, I will try to assess the most important themes and provide an adequate overview and insight into the relation prevalent between institutional investors and exchanges or the development of different exchange organizational forms, respectively. I will focus on trends and developments applying their impact on this relationship that have evolved after the crisis as well as those already predominant in the years leading up to 2007.

A. Wöhrmann (✉)
Deutsche Asset & Wealth Management Investment GmbH,
Frankfurt am Main 60612, Germany
e-mail: asoka.woehrmann@db.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_22

22.2 Structure Organization

The remainder of this chapter is organized as follows. In a first step I will briefly review institutional investors' evolution as one of the key stakeholders acting through exchanges. In particular, I will focus on the ongoing growth of institutional investors, which is supported by households increasingly delegating the decision about their savings' investment to professional investment managers. At the same time, being an asset manager, I will also address the concurrent asset management industry's consolidation, which is reflected in an ever-increasing amount of assets under management coinciding with a steadily decreasing number of asset managers. Hereafter, I will try to lay out the motives driving institutional investors to conduct their trading through exchanges. While doing so, I will introduce you to various steps and participants being involved in a trading process focusing on the importance of liquidity, transparency, and their impact on determining transparent and fair market prices.

This section is followed by an overview of recent historical and concurrent developments revolving around exchanges and their different organizational forms. In particular, I will display some of the most crucial regulatory changes concerning equity markets and subsequently analyze their effects on markets' microstructure. Both over the course of this section and thereafter I will introduce you to different sets of trading venues investors can choose from when considering execution of their trades.

Last but not least we will have a closer look at two new organizational forms, which have been growing at a notable speed over recent years.

22.3 Institutional Investors and Exchanges

Henceforth, when referring to the term "institutional investors," I will follow Davis and Steil (2001) defining institutional investors as "specialized financial institutions that manage savings collectively on behalf of other investors based on specific objectives in terms of acceptable risk, return maximization and maturity of claims."

The Bank of International Settlements [1] adopts this definition and includes insurance companies, pension funds, and investment companies being investment vehicles in their own right (e.g., mutual funds and hedge funds).

This group of investors has gained significant influence on markets' developments over the recent years by growing at a considerable speed.

In its annual report 1998, the BIS recognized and clearly stated the importance of institutional investors and predicted their significant growth path: "Institutional investors are a permanent feature of the financial landscape, and their growth will continue at a similar, and perhaps faster pace. The factors that underpin their development are far from transitory and in many cases have only just started having an impact. The behavioral characteristics of institutional investors, therefore, will be an increasingly important determinant of domestic and international financial market conditions, and the implications for financial market stability warrant serious consideration."[2] Davis and Steil (2001, Executive Summary) share this conviction and highlight the influence of households' increased tendency to rather trust their savings to professional portfolio managers instead of becoming active on security markets themselves or holding them as bank deposits. The authors describe this process as "institutionalization of savings."[3]

According to them, institutionalization is linked to various supply and demand factors including the ability to take advantage from technological developments and increased competition. Both of them will be addressed later on.

The result, however, is present in the institutional investors' growth in the past decades. Davis and Steil note that institutional investors at this time accounted for 30 % of household sector assets and 30 % of financial intermediation. In line with these numbers, the BIS sees an ongoing and rapid catching up of institutional investors' assets under management with those of the banking system and thereby a contribution to improved financial markets' functioning and depth (BIS, 2007).

While in 2003 commercial banks in total had USD 49 trillion of assets under management, institutional investors' assets under management amounted to USD 47 trillion.

Furthermore, the report confirms Davis and Steil's observation of an increasing share of households' savings being channeled through institutional investors stating a rise of household sector holdings' average of assets managed by these investors from 36 to 44 % between 1995 and 2005.

However, it is not the sheer size of institutional investors that matters for the examination of their relationship with exchanges. It is rather how this increased size might affect exchanges, e.g., with respect to trading volume, liquidity, and volatility.

The SEC[4] for example finds that overhauls that occurred to the exchanges' market structure, which I will discuss in greater detail later on, were largely driven by institutional investors' trading and led to a huge increase in trading volumes. Referring to the average daily volume on the New York Stock Exchange (NYSE) they noticed a jump of 1600 % between 1990 and 2013. The SEC further concludes that investors are "dominant market players," albeit pointing out the differing interaction with the market and subsequently differing impacts on markets among the members of the broad classification as institutional investors. However, the SEC points out that, in general, institutional investors "are known to improve price discovery, increase allocative efficiency and [...] provide trading markets with liquidity."

In line with this opinion, the BIS (2003) notes that "efficient markets [...] require the existence of investors with enough capital and sufficiently long investment horizons to arbitrage potential misalignments away so as to fully incorporate available information into prices." Although this description generally fits with institutional investors, the BIS, while having a closer look at broader trends within the asset management industry, points out some developments and tendencies that might in fact harm market efficiency.

More narrow and tiered investment mandates such as the choice of market benchmarks in order to assess performance in combination with standard compensation structures might prevent investors to trade against each other, i.e., from taking contrarian positions. This might trigger losses in their ability to eliminate mispricings and to reduce feedback trading or herding effects in the market. On the contrary, as many benchmark indices are chosen from a set of established indices, the BIS highlights an effective performance measurement against a peer universe, which in turn might rather spur feedback trading and shorten the investment horizon of institutional investors.

The BIS indicates in its paper that feedback effects might also root from institutional investors' herding behavior. Following Merli and Roger (2012), herding is broadly defined as an investor's imitation of the actions of others, i.e., other investor's investment decisions. Frey, Herbst, and Walter (2007), for example, show herding behavior on the basis of a data set of German mutual funds.

However, these are developments induced by institutional behavior on exchanges and lie beyond the scope of this chapter.

22.4 (Types of) Exchanges

Here I want to keep focusing on how the growth of institutional investors might have affected the equity market's landscape and how exchanges organize.

In order to address these topics, first we have to understand institutional investors' motivation to trade through exchanges. Naturally, even before, we have to clarify what the term "exchange" implies for us, investors, and regulators.

Henceforth, I will therefore apply the US Security and Exchange Committee's definition considering an exchange being an organization, association, or group of persons that firstly brings together the orders for securities of multiple buyers and sellers and that uses established, nondiscretionary methods (whether by providing a trading facility or by setting rules) under which such orders interact with each other, and the buyers and sellers entering such orders agree to the terms of a trade.

To put it in simple words, exchanges are markets bringing together supply and demand for securities. Thereby, they mediate between capital providers and capital requesters.

By providing a common meeting place and time, i.e., by providing traders with the ability to locate each other, exchanges free both parties from initiating direct negotiations among each other. This on the other hand, due to the intensive research required to find the (best) matching counterparty, i.e., the search for liquidity, would be very much more expensive otherwise (Angel et al., 2012).

Exchanges therefore are markets, which are designed to minimize the cost related to search and find an adequate trading partner.

Following the CFA Institute (2012) most registered exchanges have adopted the structure of an electronic limit order book market being characterized by a multilateral nature (multiple buyers and sellers can trade against each other within the system), a nondiscriminatory access, and an alignment of their operation with nondiscretionary rules and procedures.

Naes and Skjeltorp (2006) characterize the limit order book market as a market, where investors themselves provide liquidity and set prices. Following the authors a limit order can be defined as a buy or sell order for a specific volume and price determined by the buyer or seller, respectively. The investors themselves provide liquidity by placing limit orders or demand liquidity by placing market orders. The latter are orders to buy or sell at the current price in the limit order book, i.e., focusing on execution timing rather than execution price. However, an important, typical feature of such limit order books is predominant in their price-time priority, which is thought to ensure a fair treatment of orders (CFA Institute, 2012).

Trades are established in case of "acceptable" matches between buy-side and sell-side orders (CFA Institute, 2012), i.e., trades generated by electronically matching orders on the basis of predefined rules (Naes and Skjeltorp, 2006).

For this to happen, however, the spread "has to be crossed." This jargon term implies nothing else than the acceptance that either the buy side or the sell side has to accept the higher or lower price, respectively, in order to achieve overlapping orders and a subsequent execution of the trade (CFA Institute, 2012).

The CFA Institute (2012) lays out that brokers only enable the execution of client orders in limit order book markets, but also indicates the inclusion of liquidity providers in order to complement the interaction of customer orders by many exchanges. This serves the facilitation of a continuous market's operation.

Moreover, the authors note that exchanges are generally both pre-trade and post-trade transparent. This refers to the public availability of pricing and trading interest prior to execution on the one hand and public dissemination of the transaction details after the execution on the other hand. At most exchanges, pre-trade transparency is ensured by a public display of the top-of-the-book trading interest in the consolidated quote stream if it constitutes the national best bid and offer (NBBO) for a stock. The depth of the order book is displayed to the exchanges' participants. Post-trade transparency, i.e., disclosure of the transaction details, happens in real time at exchanges. This transparency allows exchanges to be commonly known as lid liquidity, albeit this is in some way not completely correct as also exchanges might have some dark liquidity parts (CFA Institute, 2012).

But we will get back to this later on, when discussing the different concepts of liquidity.

Staying with exchanges right now, looking at the NYSE, it is interesting to actually face a so-called hybrid market (Naes and Skjeltorp, 2006). This term describes that the major part of trading is channeled through the limit order book, while dealers (here known as specialists) are required to set prices for the stocks they are responsible for, if liquidity shrinks under a certain threshold.

Thus, the NYSE is a combination of the two different types of markets Naes and Skjeltorp differentiate: the limit order market and the dealer market, which presents another market form exchanges can adopt (e.g., NASDAQ).

In a dealer market groups of intermediaries are responsible for setting tradable prices in their stocks. The dealer has to provide liquidity in the form of an adequate share inventory and earns the difference of the bid-ask spread in return.

Based on information models, in particular referring to Copeland and Galai (1983) the authors note that dealers might not be able to differentiate between informed and uninformed investors. As losses would incur for dealers when trading with informed investors, dealers will always set a positive spread in order to compensate for the expected loss. However, the authors elaborate further, referencing to Glosten and Milgrom (1985) that the order flow will teach dealers and other uninformed investors about the correct price as private information can consequently be incorporated. Finally, taking another study by Easley and O'Hara (1987) and thus a strategic component into account, they conclude that early studies of the transaction process suggest the spread's composition being comprised of information costs on the one hand and inventory costs on the other hand. However, they also point out that there is a more recent group of studies, which coincided with the emergence of order-based trading models and which assumes a potential strategic behavior by liquidity providers due to market power or access to private information. Naes and Skjeltorp (2006) state that these models' main outcome consists in identifying oligopoly rents for liquidity providers with market power.

With respect to transparency, dealer markets usually possess much lower levels of transparency compared to limit order markets. Resuming the outcome of a number of theoretical studies, they point out that increased transparency culminates in better liquidity and reduced transaction costs. Nonetheless, they also reference to Madhavan's study (1995) showing that transparency is not necessarily improving liquidity, but might rather decrease liquidity when orders are withdrawn due to the investors' unwillingness to reveal their interest in buying or selling a security.

After having seen the two types of markets that exchanges traditionally were organized as, it is obvious that institutional investors' main motivation behind trading at exchanges lies within their search for liquidity and price efficiency (i.e., a correct and informative price discovery) to execute their trades.

Naes and Skjeltorp (2006) refer to this when formulating two main functions of a market, i.e., the provision of liquidity for buyers and sellers as well as to ensure that prices reflect relevant information about fundamental value.

Following O'Hara and Ye (2011), a market's ability to meet this dual goal of liquidity and price discovery is termed a market's quality.

At this point in time it might be helpful to have a closer look at two things, liquidity on the one hand and how trades are executed in general on the other hand.

22.5 Participants in a Trade and Transaction Costs

A trade originates from the interaction between the buy side and the sell side. According to Harris (2010), institutional investors would represent the buy side, whereas brokers and dealers would be attributed to the sell side. Dealer and broker sell their trading services and thus both exist in order to facilitate the execution of the buy side's desired trades. Obviously, their services are not for free, albeit the rewarding for the trade's execution differs according to the different approaches used by dealers and by brokers.

Dealers directly trade with their clients. When trading a security, they enter a position and take over their clients' trading problem. This is, when their clients want to sell a security, they first buy the security themselves before searching a potential counterparty and then reselling the security to them. For taking this position and thus the risk being attached (e.g., not finding a counterparty buying for the same or a higher price) they charge a spread, the so-called bid-ask spread.

Brokers, on the other hand, rather act as an intermediary. When a client places an order with them, instead of taking a position and actively trading the security they link different orders and execute the trade between the two counterparties. For a broker's services the client will be charged a commission.

Keim and Madhavan (1998) raise the idea of considering the total transaction costs as two tiered: according to them they are comprised of an explicit and an implicit component. The authors primarily consider the brokers' commissions as the explicit share of the transaction costs while including the spread, opportunity costs associated with not getting to trade at the desired time, and the possible price, i.e., market impact as a result of a trade, which will take a main part in our discussion about the relationship between exchanges and institutional investors.

In fact, the market impact is an expression of a market's liquidity. Thus I would like to have a closer look at liquidity here, which is "the lifeblood of our capital markets," as SEC Commissioner Luis A. Aguilar puts it in his speech.[4]

However, the BIS (1999) highlights that market liquidity is rather an elusive concept, for which it is difficult to find accurate and cohesive definitions.

Based on their review of single research papers they postulate a definition, which reflects the main concern when discussing liquidity of a market: a liquid market is a market where participants can rapidly execute large-volume transactions with a small impact on prices.

As to measure liquidity I will follow them denoting three dimensions of market liquidity:

Firstly, one can assess the tightness of prices, which is defined as how far transaction prices diverge from the mid-market price, i.e., the general costs incurred irrespective of market prices' level. Normally, this dimension can be measured by the prevailing bid-ask spread.

Yet, the bid-ask spread can be measured in several ways, each differing slightly in their interpretation. Following the BIS (1999) the quoted spread is defined as the gap between quoted bid and ask prices, and is preserved ex ante. The realized spread on the other hand refers to the gap between weighted averages of the bid and ask prices for executed trades over a period of time. The transaction volumes at each price fix the weights. Another instrument is the effective spread relying on the actual transaction price, which is able to indicate the direction of price movements due to including the difference between the quoted and the actually executed price.

The second dimension refers to the so-called depth of the market, which indicates either the volume tradable in the market without price affection or the extent of market makers' order books. Measures of depth are related to the amount of trading that can be absorbed before a price divergence is caused by an imbalance in the order book. The more imbalanced the orders are, the further the market price has to diverge from the standard bid or ask price for the imbalance to be cleared. This is to say that depth in this case is measured by the fluctuation in quotes or bid-ask spreads resulting from actual order execution or, even more accurately, by also including potential trading needs stemming from portfolio adjustments. These price fluctuations are referred to by the term "market impact," although sometimes this term also includes a market's resiliency.

Resiliency (the third dimension) means the dissolving speed of potential price fluctuation caused by trades. It also can refer to the pace, with which imbalances in order flows are adjusted.

However, the applicability of measures might vary across markets and some of the measures might diverge in directions.

The three dimensions can be affected by various factors. With regard to the aim of this chapter, let me focus on the effects imposed by a market's microstructure.

Following the BIS (1999) a market's microstructure among others comprises trade execution systems, trading commissions, disclosure of contracted price, and volume information as well as market regulations.

We have already had a look on the two broad definitions of trade execution systems—dealer markets, which are also referred to as "quote-driven" markets, and auction-agency markets (e.g., limit order book markets).

The BIS also highlights that order-driven markets, i.e., auction-agency markets, are perceived to provide more efficient price discovery, that is, available information are better embedded in current prices. Yet, dealer markets are supposed to provide greater immediacy, i.e., a faster trade execution at given prices.

Another difference, as I also already pointed out, is incorporated in the differing information dissemination to market participants. While quote-driven markets provide dealers with a monopoly over order-flow information and information available to the wider market is thereby reduced, order-driven markets make more order-flow information publicly available.

Another factor linked to a market's microstructure worth examining is a market's transparency, which the BIS defines as market participants' ability to observe information in the trading process. They point to O'Hara (1995), who explains that "information" in this context is usually categorized as either public (available to all market participants) or private (not available to all market participants and including both "inside" information about fundamentals and information on order flow). According to the BIS the theoretical literature suggests that informed traders might profit from lower transparency and traders tend to delay their transactions in the hope for further information. Referencing to Scalia and Vacca (…) they note that both hypotheses are supported by empirical evidence. It is again important to be aware that increasing transparency is not always beneficial, though.

Now, I already scratched another topic being related to liquidity and of crucial importance for investors: trading costs.

The BIS (1999) confirms the separation we already saw conducted by Keim and Madhavan (1998). It further notes that implicit trading costs often impose a trade-off. This roots from a more dimensional problem investors face. On the one hand, naturally, the price of a security is decisive, whereas potentially conflicting with the time at which the trade is desired. Thus, the investor in this case might have to trade off between the costs incurred from an unfavorable price and different entry time into the position due to delaying the trade execution.

In fact, there is another dimension incorporated, which might and might not be clearly visible within the price of the executed trade—the information leakage a trader has to accept when executing the trade. This is an important point, which I will not lose of sight, but leave it for now.

22.6 Why Institutional Investors Trade at Exchanges

Let me take a step back here and point out what is important to remember.

Before starting to discuss the overhauls that have taken place in the equity markets around the world and especially in the USA as well as their implications, we have to be aware of what really drove the changes.

As a matter of fact, after having pointed out the growing size and influence of institutional investors on equity markets, after having introduced you to the general function of an exchange, and after finally highlighting the various dimensions of total trading costs, we are now able to address the driving force.

The transition in equity markets was largely driven by investors' demand for better solutions to past and prevailing trading problems (Angel et al., 2010).

The authors continue to identify the main trading problems, which, according to them, are likely to remain unchanged. However, basically all problems root back to a single challenge investors face in competing with other investors and while trying to deliver performance. This challenge consists in the investors' striving to minimize transaction costs.

The authors hint towards another crucial interest that institutional investors manifest and I slightly touched before. Large institutions, especially against the background of investors' increasing amount of assets under management, might assume large positions—also in the trades they execute. The authors hereby refer to so-called block trades.

Following Harris (2012), these orders are defined by being too large to fill easily using standard trading procedures.

And yes, there were and still are some issues related to trading larger positions.

First of all, large traders face concerns about informed trading (Angel et al., 2010). Those concerns are reflected in most traders' presumption that well-informed traders originate large trades and they might consequently face losses when agreeing to the trade. This, however, results in higher prices for such trades stemming from the risks imposed by committing to them.

On the other hand, other traders might increase the implicit trading costs for large traders by either front-running their marketable orders or by employing so-called quote-matching strategies to extract option values from their standing orders (Angel et al., 2010).

Front-running is triggered by a trader's exposure of interest to trade. Although the intention behind the exposure is to help other traders locating and subsequently filling the interest, this move might trigger undesirable consequences, as the authors raise.

Due to the previously described concerns regarding large trades, substantial averse price movements have to be accepted to encourage other traders to endorse the trade and fill the large order. These price concessions increase with the others' perception of trading with an informed trader.

Now, front-running is based on the expectations of these price changes. As large buy orders are likely to pull prices up, other traders, who became aware of the buying interest, may immediately buy in front of the order trying to benefit and generate profit from the expected price change. Naturally, this is also true with sell orders in the opposite direction. Filling orders might consequently be much more problematic and ultimately much more expensive.

On the other hand, traders who posted limit orders or quotes and become aware of the potentially trading with large traders will replace their orders and quotes in order not to lose through the negative price pressure imposed by large traders. If they succeed to take the orders from the market, the large traders will end up with higher expenses for filling their rates.

However, although these problems still persist, albeit in another form, we have to get rid of one picture—the one of institutional investors planning and executing this one big block trade.

And thereby the stage is prepared to have a closer look at what happened over the last few years with regard to equity markets. We will address various different influencing factors—technological advance and regulatory changes are usually named first when examining the research hereto—and their respective implications.

Let me begin with a topic that everybody of us has experienced and which therefore, naturally, also affected trading. Over the last decade technology has advanced with great strides facilitating traders with many new opportunities. Intermediation in trading processes changed significantly (Gomber and Gsell, 2006); costs for the search of liquidity as described before have been greatly reduced by computing technologies (Angel et al., 2010). In explaining the overhaul, which has taken place through technology change, I will follow Angel et al. (2010), Gomber and Gsell (2006), as well as CFA Institute (2012).

It appears useful to recall the process by which trading takes place for institutional investors. Usually the investor originating the order, the buy side, transmits its order to its broker/dealer to handle routing, i.e., the search for liquidity and an adequate counterparty, and execution of the order if the investor does not run its own system for doing so (CFA Institute, 2012). Remember that at the beginning of equity trading, a trade only could be executed on a floor of an exchange. Computing technologies in that sense are the continued trend to provide traders with remote access (Angel et al., 2010), which began being implemented through telegraphs being followed by telephones.

However, intermediation in trading processes has undergone significant changes based on technological developments. Brokers developed a new business model, where orders were basically transmitted directly to the markets. This model called “direct market access” (DMA) reduced trading costs as it is offered for a lower fee than traditional services and provided institutional investors with the opportunity and ability to set up their own execution concepts. Based on DMA some recent concepts like algorithmic trading, smart order routing, and liquidity aggregation emerged (Gomber and Gesll, 2006).

In fact, this brings us back to the picture of an investor placing this one, huge order to the system, which I wanted to question.

Trading today is different and the majority of institutional orders are filled using computerized algorithms, which break up the initial order (the “parent” order), slice it into smaller pieces (“child” orders), and distribute them into the market aiming at a minimization of the price impact and thereby reducing implicit trading costs. Algorithms define an order’s parameters as of what, when, where, how much, etc. is traded while minimizing market impact and information leakage, weighing immediacy of execution versus execution costs, etc. for the most efficient execution (CFA Institute, 2012). Thus, as Gomber and Gsell (2006) put it, algorithmic trading presents a broker’s competence of slicing a big order into a multiplicity of smaller orders and of timing these orders to minimize market impact via electronic means. Smart order routing concepts screen all trading venues potentially available for filling an order. Then the concepts forward the order automatically to the “best” venue while considering the different attributes of each. Such algorithms may be “aggressive,” for example, and seek to take liquidity quickly at many different trading centers, or they may be “passive,” and submit resting orders at one or more trading centers and await executions at favorable prices. “To the extent they help customers cope with the dispersal of liquidity among a large number of trading centers of different types and achieve the best execution of their customers’ orders, the routing services of brokers can contribute to the broader policy goal of promoting efficient markets” (SEC, 2010). The importance of algorithms and smart order routing is steadily increasing with the institutions’ growing size and the progressing fragmentation of equity markets—a topic that we will focus later on.

The authors hint that the responsibility of order execution quality was transferred from brokers to the buy-side trading desks with the advent of these new concepts. Indeed, this is clearly noticeable in our daily business. The amount of data that is used in these algorithmic trading processes and that is transferred through the smart order routing is simply too much to constantly monitor. Naturally, at the end of the execution process, we see time, location, and price at which the order was filled and the trade was executed, but what has happened in between origination and final execution of our order is indefinable for us. However, it is important for us to keep close track of our orders and ensure the most favorable transaction for our clients. Hence, we have to embrace this challenge.

However, technology advance has further impact on trading processes. It allowed new, powerful players to enter the market—so-called high-frequency traders (HFTs). Optimizing the speed with which their algorithms work and minimizing latency, i.e., the time it takes to observe the market event through the time it takes to analyze this event and send it back to the venue that responds to the event (Hasbrouck and Saar, 2010), they assume arbitrage strategies to exploit temporary mispricing in the markets. The most frequently used definition of HFT comes from the SEC (SEC Concept Release on Equity Market Structure, 75 Fed. Reg. 3603, January 21, 2010). They are defined as "professional traders acting in a proprietary capacity that engage in strategies that generate a large number of trades on a daily basis." They are characterized by (1) use of extraordinarily high-speed and sophisticated computer programs for generating, routing, and executing orders; (2) use of co-location services[5] and individual data feeds offered by exchanges and others to minimize network and other types of latencies; (3) very short time frames for establishing and liquidating positions; (4) submission of numerous orders that are cancelled shortly after submission; and (5) ending the trading day a position as close as possible to flat (i.e., not carrying significant unhedged positions overnight). To put this into context, HFTs have average holding periods of a few seconds and measure speed in micro- (and sometimes nano-) seconds.[6]

The HFTs' business model also bases in the adoption of a "maker-taker" pricing model by many registered exchanges in the USA, which were implemented in order to attract liquidity providers (SEC, 2010). Hereby non-marketable resting orders, i.e., orders providing liquidity at a particular price, receive a liquidity rebate if they are executed, while those "taking" liquidity are charged an access fee. HFTs take advantage hereof by submitting large numbers of non-marketable orders using low-latency systems (often cancelling a very high percentage of them), which are supported by highly automated exchange systems (SEC, 2010).

Although the public opinion regarding HFTs seems to have a rather negative connotation, especially after the Nasdaq Flash Crash in 2010, it is important to notice that HFTs might have contributed something to a more liquid environment.

As I pointed out already, the equity market's landscape has undergone quite a few changes over the last decade so that it is difficult to attribute findings to one particular change. However, the majority of empirical studies point to a reduction in transaction costs, explicit and implicit, since the computerization of equity markets started.

Jones (2013) delivers a very comprehensive overview over the research done so far on this topic and draws a rather positive bottom line: market liquidity improved, trading costs reduced, and stock prices became more efficient. He addresses the regulator: "Minor regulatory tweaks may be in order, but those formulating policy should be especially careful not to reverse the liquidity improvements of the last 20 years." Indeed, I think that HFTs might act as new market makers, hence benefitting investors by providing liquidity. This is in line with Angel et al. (2013), who state: "Public investors benefit when high frequency traders offer liquidity when dealing, and move liquidity among markets by arbitraging markets." In fact, we need those arbitrage strategies in order to uphold a continuous market. If we think of base-level economic lessons, arbitrage processes lead to a correct price finding. Therefore, HFTs might contribute positively to markets' functioning. Nevertheless, they also apply some "predatory" trading strategies, which might harm investors by increasing adverse selection[7] and subsequently investors' transaction costs. NBIM (2013) gives a nice overview of such strategies, but I want to highlight especially one of them, which is very important to institutional investors—order anticipation, the modern front-running. As Angel et al. (2013) note, transaction costs, albeit having shrunk significantly, might still be elevated by traders front-running institutional investors' large trades using sophisticated tools, which detect trading patterns and take advantage of these.

However, from my point of view, the initial advantages of HFTs, which were notably present some years ago, are decreasing steadily. Speed, one of the key aspects to their arbitrage strategies, is becoming increasingly unimportant and it is questionable whether the sums being spent on higher velocity are still justified. From the standpoint of an institutional investor it is definitely important to note that our investment horizon normally stretches out multiple years. Of course, we have a certain responsibility towards our clients to execute the order most efficiently possible. But we have to notice that price reactions taking place in the time intervals measured nowadays will not or only slightly affect the total return over the longer period.

Let me now turn our focus to a very much discussed and clearly visible development currently reshaping the equity market's landscape—literally.

Markets' fragmentation is not a new phenomenon. Nevertheless, it only has strongly gained importance recently.

As Greese (2012) points out, exchanges have been and sometimes still are perceived as natural monopolies due to a concept, which is called the "virtuous circle of liquidity" and is based on Mendelson (1987). According to this concept, traders search for the market offering the best liquidity, which, following Harris (2011), is due to the fact that it allows them to cheaply implement their trading strategies. Greese (2007) referencing Mendelson (1987) reasons forward that the most liquid market would be the one presenting the most participants due to the provision of the highest probability of order execution and the most competitive prices. With reference to Pagano (1989) he concludes that the market with the highest number of traders should attract all other participants, leading to a consolidation of order flow in a single market due to "liquidity begetting liquidity."

Indeed, looking back, such behavior was visible in equity markets. As Angel et al. (2010) note, the advance of telecommunications technologies firstly led to a consolidation of the exchange industry. The authors highlight that consolidated quote feeds mandated by the SEC and sold by various data venders enabled traders to be aware of all order sizes at the best bit and offer. Thus, it was easy for them to identify the markets offering the current best trading opportunities. Investors sent their orders to larger markets expecting the highest probability to find the respective trading partner. Since both trading parties could find each other more easily, transaction costs decreased while trading volumes increased. Order flows consolidated and exchanges such as the NYSE obtained up to 90 % market share.

However, Gresse (2012) denies the current existence of such consolidation and is clearly supported by literature, which highlights once and again that order flow is fragmented, both in the USA and in Europe. O'Hara and Ye (2011) for example observe that nowadays almost 30 % of US volume is being traded off the primary exchanges.

So we have to discuss what happened in between.

If we take a look at the global map of equity trading, although to a different extent, fragmentation is visible both in the USA and in Europe, but not so much in Asia for example. This is due to the regulatory adjustment that has taken place on those two continents, in particular the implementation of the Regulation National Market Systems (Regulation NMS) and the Markets in Financial Instruments Directive (MiFID) in the USA and in Europe in 2007, respectively. These regulatory amendments were attempts to catch up with the technological progress over the years before and spurred both the competition between traditional exchanges and the accelerating so-called alternative trading systems (ATSs).

Let me begin with the USA. In 1975 the SEC introduced the National Market Systems (NMS) with its core component the Intermarket Trading Systems (ITS). The latter links all US markets trading exchange-listed securities except NASDAQ securities (Gomber and Gsell, 2006). When implementing the ITS Plan (SEC, 1978) the SEC urged "best execution," i.e., routing the orders to the market offering the best price generally granting the receiving markets 30 s to respond (SEC, 2005). As Gomber and Gsell (2006) note, this is set forth in the so-called "Trade-Through Provision" in the ITS Plan. They define a trade-through by the execution of an order despite the availability of a better price at another market. As the SEC (2013) states, these intermarket rules protecting displayed quotations against trade-throughs and locking/crossing quotations applied to both automated quotations of electronic venues and the much slower manual quotations of floor-based exchanges. The SEC continues to show that Regulation NMS adopts new rules applying uniformly to all US-listed stocks and only protecting automated quotations[8] (order protection rule). Its objective hereby was to promote fair competition, i.e., abolish the competitive advantages slower manual markets possessed due to former rules.

In a nutshell, the Regulation NMS's key aspects include the Access Rule, which ensures fair and nondiscriminatory access to markets and prices for market participants, the order protection rule, and market data rules, which set the framework for allocating revenues to centers contributing data to the consolidated quote and tape (CFA Institute, 2012). The SEC (2013) notes the affection by Regulation NMS and its fragmenting effects. According to them, the NYSE's market share in its listings declined from 79 % in 2005 to 25 % in 2009, while the total volume in NYSE-listed stocks during this period increased by 181 % with the introduction of more automated trading on the NYSE and elsewhere (SEC 2010).

The SEC Concept release (SEC 2010) displays very well how Regulation NMS affected the equity market structure. Order flow is fragmented among four types of trade execution venues now.

Firstly, registered exchanges: In 2010 approximately 63.8 % of share volume in NMS stocks was executed by registered exchanges, which characterize themselves by their self-regulatory responsibility. We already mentioned the high decline in NYSE's market share, which we can put into context that no single exchange executes more than 19.4 % of share volume in NMS stocks. Hence, the SEC has achieved its objective to foster competition among trading centers.

So-called electronic communications networks (ECNs) present a second trading center form competing in the current market structure. Following the CFA Institute (2012) ECNs are multilateral electronic trading venues operating similarly to exchanges in terms of secondary market trading of equity securities. Typically they are structured as limit order book markets and generally are pre- and post-trade transparent, albeit being considered an ATS, which are not required to publicly display price quotations and which are able to restrict access to their crossing systems and internalization pools. According to Angel et al. (2010) ECNs at the beginning were not very successful in competing with exchanges as too much order information remained on the floor and there were more trading opportunities available. Nowadays most ECNs that competed with the primary exchanges have since been acquired or merged with the large exchange operators (CFA Institute, 2012).

A third and very important form of trading center is so-called dark pools. Following the CFA Institute (2012) dark pools are defined as "systematized execution facilities that operate without full pre-trade transparency." Thus, contrary to ECNs, orders entered in these ATS are not displayed to other market participants and matched anonymously against counter-side orders, i.e., their best-priced orders are not transmitted for inclusion in the consolidated quotation data (SEC, 2010). For such venues the terms "dark venues" or "dark liquidity," including dark pools and broker/dealer internalizations, as opposed to lid venues or lid liquidity, including exchanges and ECNs, were coined (Zhu, 2012).

As a matter of fact, the concept of dark liquidity is nothing particularly new to the world of equity trading. And again the main driver perhaps is the interest of institutional investors to minimize their trading costs, in particular the implicit costs incurring from adverse price movements triggered by large trades. As the SEC (2009) notes with reference to older releases that "large investors often seek ways to interact with order flow and participate in price competition without submitting a limit order that would display the full extent of their trading interest." It then continues to present some of the ways investors can accomplish this aim. One of them is "to use a trading mechanism that permits some form of "hidden" interest to interact with the other side of the market."

Institutional investors previously achieved their aim by arranging their orders with so-called block traders, i.e., block broker or block dealer. Those would either try to find the adequate counterparty or fill their clients' large orders. Another way would have been to place small slices of the block trade in order to prevent displaying the signal of a large block trade coming up. Today, algorithms and dark liquidity venues take over these tasks to a large extent. Algorithms may often cancel their orders while posting liquidity in order to cover their presence and thus to prevent other traders from exploiting information incorporated in their orders. This is most commonly done by immediate-or-cancel orders (IOC) (Angel et al., 2010). These tactics are feasible thanks to the strong decline in latency (NYSE's average speed of execution for small, immediately executable orders declined from 10.1 s in 2005 to 0.7 s in 2009[9]). Dark liquidity venues, which can have different designs regarding the services they provide for their clients, from my point of view grant a more trustworthy system than block traders and block dealers did. These type of market makers had in the past much more time to act and position themselves or other parties before executing the orders, effectively making it more expensive for institutional investors to trade. And these newer venues continue to gain market share. According to Zhu (2012) with reference to data from Tabb Group and Rosenblatt Securities, the market share of dark pools roughly doubled from around 6.5 % in 2008 to about 12 % in 2011 displaying the institutional investors' increased need for possibilities to fill their orders without exposing their interest to the broader market as the order sizes and depths on exchanges have declined dramatically (Chordia, Roll, and Subrahmanyam, 2011).

However, before assessing the various forms of dark pools, I should stress that there also exist a fourth form of trading facilities at the ECNs and exchanges usually classified as lid venues that provide investors with dark liquidity. These facilities are called hidden-order facilities and allow traders to submit orders that limit the exposure of their sizes (Angel et al., 2010). There are different order types deviating in their degree of information exposure: hidden orders completely omit the true size of the total order, while reserve orders partially reveal size and discretionary orders reveal size in whole or part at prices away from the market. Again, the aim is to prevent information leakage and the potentially adjacent front-running or quote matching. The hidden liquidity is discovered through submitting orders to trade at that price. However, the price is a binding commitment to trade with the hidden size.

Apart from these facilities, as stated before there are various other forms of dark liquidity. Closely following the CFA Institute (2012) let me first assess the different types of dark pools operating in the USA. They refer to categories established by Rosenblatt Securities and TABB Group. While the latter categorizes dark pools according to their function (block-cross platforms, continuous-cross platforms, and liquidity-provider platforms), Rosenblatt Securities rather orientates along the group operating the pool (pools operated by bulge-bracket brokerage firms, pools operated by market makers, independent or agency pools, and consortium-sponsored pools).

Adopting these classifications, independent/agency and consortium-sponsored dark pools are mostly black-cross platforms, meaning that trades take place in periodic auctions within the system, in which the volumes crossed at each point are of large size. On the other hand, broker/dealer-operated dark pools are most commonly continuous cross platforms. Here crossings take place more frequently and usually incur a smaller size.

Additionally one can differentiate between dark pools being crossing networks, dark pools only facilitating the matching of customer-to-customer order flow, and others allowing customer order flow to also execute against the broker's own account.

The CFA Institute (2012) defines crossing systems as automated systems that match order flow in an orderly or systematized manner between counterparties affiliated to the same system or network.

As with regard to the pricing, orders in these crossing systems are typically crossed at a point within the spread of the best bid and offer reference prices. Thereby, those dark pools can provide limited priced discovery (Zhu, 2012).

This brings us to discuss the investors' benefits from using dark pools. We already discussed the prevention of information leakage. Apart from that, dark pools can offer investors a price improvement as well as reduced transaction costs.

Mostly dark pools derive execution prices from lid venues. For example, a typical, classical form of dark pools matches customer orders at prices derived from lid venue, such as the midpoint of the national best bid and offer (NBBO) or the volume-weighted average price. Thereby, dark pools offer investors a price improvement with regard to an exchange for example as they save on the bid-offer spread (CFA Institute, 2012). However, taking lit-venue prices as reference prices, these dark pools do not provide direct price discovery.

Also, the order submission to dark pools, due to the absence of market makers, is subject to a trade-off between potential price improvements as described above and the risk of failing execution (Zhu, 2012).

On the other hand, broker/dealers might have an incentive to set up their own dark pools enabling them to better match orders internally. This in turn frees them from trading fees they had paid to exchanges and other trading centers otherwise (Zhu, 2012) benefitting investors in terms of (potentially) lower trading costs not only from these dark pools, but also due to decreased access fees for exchanges as they compete with the dark pools for order flows.

Additionally, some dark pools conduct liquidity aggregation; that is, they endow investors to use liquidity across different sources. Thereby, available liquidity is increased and attractiveness for executing large orders is heightened. Finally, dark pools might benefit institutional investors by restricting some market participants, or more precisely their predatory trading strategies, such as high-frequency trading firms. However, recently HFTs have selectively also been granted access to dark pools and increasingly adopt the role as market makers there, too. This is beneficial as it endows investors with the ability to cover more dark pools at once. Protection against predatory trading strategies can be achieved by posting minimal average quantities (MAQs). As HFTs use to "ping" the market probing for hidden trading interest with small quantities, MAQs prevent investors' large orders from being discovered.

Another type of dark liquidity is the so-called internalization. The CFA Institute (2012) denotes internalization as the process of brokers/dealers executing client order flow against their own accounts on a systematic basis. Hence, trades are executed on a bilateral basis and the broker/dealer assumes the role of the counterparty to all incoming orders, trading as principal and using its own risk capital. Like dark pools and ECNs it is therefore referred to as off-exchange activity. As over-the-counter (OTC) market makers (broker/dealer) are not required to post quotes prior to execution, internalization also is included in the term "dark liquidity" (CFA Institute, 2012).

As Zhu (2012) elaborates very well, there are quite a few concerns regarding market efficiency and price discovery related to dark liquidity and dark pools in particular.

The SEC (2013) conducted a literature review regarding the effect of dark fragmentation, i.e., the dispersal of volume from lid venues to dark venues. Let me briefly summarize the key empirical aspects of some of the reviewed papers as elaborated by the SEC.

Boni, Brown, and Leach (2013) find that dark pools, which are specifically designed to foster buy-side exclusivity, show some evidence for smaller execution footprints, i.e., market impact, and higher overall execution quality for large trades. Buti, Rindi, and Werner (2011) note a joint determination of market quality measures and dark pool activity. Their findings suggest better market quality being associated with more dark pool activity. According to them this provides narrower spreads and more market depth. On the other hand, they find a negative correlation between dark pool activity and short-term volatility, indicating improved price efficiency. Yet, other findings suggest that more dark pool activity is linked to less efficient prices for NASDAQ-listed stocks as well as for small- and medium-sized stocks. The CFA Institute (2012) find initially narrowing bid-offer spreads and increased best-price depth, i.e., improved market quality to come along with increased dark trading. However, the authors also note a threshold effect implying some turning points in the impact on market quality, i.e., increasing quality before and worsening after the turning point, depending on a stock's market capitalization. Hatheway, Kwan, and Zheng observe higher transaction costs associated with non-block dark trading, whereas dark trading in large size (defined as the top 1 % of trades by trade value in a stock) apparently implicates narrower effective spreads. Also, when not controlling for the level of informed trading, their findings suggest narrower effective spreads related to dark trading. Finally, O'Hara and Ye (2011), after controlling for number of trades, trade size, price inverse, daily returns, and market capitalization, note that higher levels of off-exchange trading imply lower effective spread.

To sum it all up, the literature notes that dark trading might harm market quality. However, there is some evidence that with respect to trading large orders dark trading does not harm but rather improve market quality.

Another concern is, however, raised by Angel et al. (2013): the authors note that exchanges' market share has decreased and concerns have been uttered regarding the widened NBBO spread and it being less informative. They do express some concern about a potential "degradation of the NBBO" while hinting to bid-ask spreads at historically low levels on the one side and very high levels of depth on the other side.

Nevertheless, they also give rise to another concern. For being regulated as broker/dealers and not as exchanges, dark venues have much lower regulatory burdens than exchanges. Interestingly, however, the authors suggest to rather lower the regulatory standards of exchanges than increasing those of dark venues.

Finally, Degryse, van Achter, and Wuyts (2008) also hint for an issue arising from the price discovery process in dark venues, especially in crossing networks (CN). As I denoted before, there is no provision of price discovery by dark venues. Trades are executed with prices adopted from existing primary markets ("parasite pricing"), which according to the authors incorporates some latent risks as it requires a sufficiently informative and well-functioning existing primary market in order to provide network users with adequate and correct pricing.

22.7 Wrap-Up: Institutional Equity Investors and Exchanges

Due to the ever-increasing assets under management with large institutional investors, the challenge to not only maintain but even enhance efficiency in executing trades becomes more important. We require best service for multiple facets of trading, with transparency and liquidity being the most important two.

Based on improved and faster technical solutions, trade execution and relevant platforms have changed substantially. To secure best execution for our clients as part of our fiduciary duty we have to make sure that we minimize potential damage from the marketplace exploiting information about our trading strategies. Finding sufficient liquidity and minimizing slippage for the entire order size bring us to engage with the market through different trading venues and with multiple trading partners.

As long as regulated and transparent exchange places provide us with state-of-the-art execution services and highest standards with respect to bringing together buyers and sellers while ensuring minimal disclosure of order information, these venues will remain an important partner in conducting our business. However, some orders will require specific handling and alternative forms of order execution will remain important and maintain pressure on established execution service providers to enhance and improve services also going forward.

References

1. BIS. (2007). Institutional Investors, Global Savings and Assets Allocation. *CGFS Papers, No. 27*. Bank for International Settlements, Basel

Chapter 23
Equity Market Fragmentation in the Swiss Market

Andreas Grünbichler, Alexander Kohler, and Rico von Wyss

23.1 Introduction

The Markets in Financial Instruments Directive (MiFID, EC [1]) and its 2008 amendment [10] changed the landscape for European Stock Exchanges. Before the directive was introduced, a concentration rule was in place that allowed member countries to require the execution of certain orders at a regulated market. This concentration rule was beneficial for established national exchanges.

By removing the concentration rule, MiFID enabled the competition among trading venues, which led to an increased emergence of alternative trading platforms and their gain of market share. A similar development took place in the USA over the decade before, where electronic communication networks (ECNs) like Archipelago, Island, and Instinet could increase their market share in the trading of US stocks on the cost of established exchanges like NYSE and NASDAQ. As a consequence, a consolidation on the level of exchanges took place with the purchase of the ECN Island by Instinet in 2002, the merger of NYSE with Archipelago to the NYSE Group, and the purchase of Instinet by NASDAQ in 2005.

In the course of the implementation of MiFID several multilateral trading facilities (MTFs) were launched in Europe, starting in March 2007 with Chi-X, a pan-European MTF owned by a consortium of global financial institutions. In 2008 several MTFs followed like BATS Europe and Nasdaq OMX Europe, two European subsidiaries of American exchanges, and Turquoise, an MTF owned by nine

A. Grünbichler • A. Kohler
Wüstenrot Gruppe, Alpenstraße 70, 5033 Salzburg, Austria
e-mail: andreas.gruenbichler@wuestenrot.at

A. Kohler
Deloitte AG, General-Guisan-Quai 38, 8022 Zürich, Switzerland

R. von Wyss (✉)
Bank Vontobel AG, Gotthardstrasse 43, 8022 Zürich, Switzerland
e-mail: rico.vonwyss@vontobel.ch

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_23

investment banks. The increasing number of trading platforms and the possibility, as well as the pressure, to choose the most efficient trading channel led to a fragmentation of trading volume. In June 2010 more than 25 % of the overall trading volume for European equities was traded on four MTFs.

Swiss stocks encountered the same development without regulatory pressure. According to the Fidessa Fragmentation Index (see fragmentation.fidessa.com) about 75 % of the aggregated trading volume of the SMI stocks in June 2008 was traded on the Swiss exchange and about 1.3 % on Chi-X (the rest was traded on dark venues, OTC, and through systematic internalizers). In June 2009 the share of the Swiss exchange had dropped to 65 % and in June 2010 to 51 % of the overall trading volume. The share of Chi-X has risen to almost 13 % and other MTFs could increase their market share as well (BATS Europe accounts for almost 6 %, Turquoise for almost 3 %, Nyse Arca and Nasdaq Europe together for 0.5 %).

In this chapter, we analyze two questions around liquidity fragmentation in Europe and its impact on equity trading in Switzerland:

1. First, how is market quality affected by fragmentations
2. Second, where does the information processing take place

23.2 Data

We conduct our analysis for the constituents of the SMI Expanded index that includes the 50 largest Swiss stocks. Stocks that were not traded on the three MTFs Chi-X, BATS Europe, and Turquoise and stocks where data was not available are excluded. Our final sample in Table 23.1 consists of 29 stocks.

We obtain intraday trade and quote data from Thomson Reuters Tick History for the Swiss exchange and the MTFs Chi-X, BATS Europe, and Turquoise. Our sample covers 433 trading days (20 months) between November 3, 2008, and June 30, 2010. The data covers trades executed in the limit order book of the Swiss exchange and the three MTFs Chi-X, BATS Europe, and Turquoise, but it does not include trades executed by systematic internalizers, dark pools, or OTC venues.

For our analysis we built one-second snapshots of historical order books containing the best bid and ask price and the corresponding volumes. Historical trade data is aggregated to one-second intervals by summing up trading volume and calculating the volume-weighted average price. Historical trade and quote data is calculated for every stock on every trading venue from 09:00:00 (CET) until 17:15:00 (CET) on each trading day.

23.3 Fragmentation and Market Quality

Table 23.2 shows the fragmentation of liquidity. Panel A presents the market share in terms of average daily trading volume, and Panel B in terms of average daily number of trades for the four trading venues.

Table 23.1 Sample

Company	Symbol	MCAP	Subsample
Nestle	NESN	168.1	Stocks L
Novartis	NOVN	135.2	Avg. MCAP: 66.0
Roche	ROG	113.8	
Credit Suisse	CSGN	53.2	
UBS	UBSN	51.3	
ABB	ABBN	43.0	
Zurich Financial Services	ZURN	32.4	
Syngenta	SYNN	24.1	
Holcim	HOLN	19.9	
Swisscom	SCMN	18.8	
Swiss Re	RUKN	15.2	Stocks M
Synthes	SYST	15.1	Avg. MCAP: 10.2
Richemont	CFR	14.9	
Kuehne + Nagel	KNIN	10.6	
SGS	SGSN	10.2	
Adecco	ADEN	9.2	
Swatch Group I	UHR	6.9	
Actelion	ATLN	6.9	
Givaudan	GIVN	6.6	
Geberit	GEBN	6.3	
Swatch Group N	UHRN	5.3	Stocks S
Lonza	LONN	4.9	Avg. MCAP: 3.2
Baloise	BALN	4.2	
Swiss Life Holding	SLHN	3.5	
Nobel Biocare	NOBN	3.2	
Logitech	LOGN	3.1	
Clariant	CLN	2.1	
Petroplus	PPHN	1.6	
OC Oerlikon	OERL	0.8	

The final sample of 29 companies consists of the constituents of the SMI Expanded index that are listed on the MTFs Chi-X, BATS Europe, and Turquoise. We use the index constituents as on June 15, 2010. Additionally, we show the attribution of the stocks to the subsamples. It is based on the average daily market capitalization (MCAP) over the sample period and is reported in billion Swiss francs. Source: Kohler and von Wyss [8]

The Swiss exchange as the traditional and established market attracts the highest fraction in terms of trading volume (80.86 %) as well as the number of trades (72.34 %). The three MTFs exhibit a substantially lower market share in average daily trading volume with 11.49 % for Chi-X, 5.01 % for Turquoise, and 2.64 % for BATS Europe and in the average daily number of trades with 16.29 % for Chi-X, 6.98 % for Turquoise, and 4.40 % for BATS Europe. As the market share in terms of

Table 23.2 Market shares

Panel A: average daily trading volume (in 1000)

		Total		SIX		BS			CHI			TQ		
		Volume	FI	Volume	Fraction (%)	Volume	Fraction (%)	Sig.	Volume	Fraction (%)	Sig.	Volume	Fraction (%)	Sig.
Pooled sample	Mean	2,914,066	1.52	2,336,969	80.86	85,733	2.64	***	350,284	11.49	***	141,080	5.01	***
	Std			113,709		8510			22,574			8609		
Stocks L	Mean	2,330,061	1.55	1,846,450	79.85	73,229	2.83	***	290,672	11.99	***	119,710	5.33	***
	Std			140,518		13,273			31,466			11,480		
Stocks M	Mean	421,388	1.44	346,787	83.16	10,344	2.17	***	46,948	10.51	***	I7,309	4.16	***
	Std			29,167		2058			6210			2166		
Stocks S	Mean	162,617	1.26	143,731	89.06	2160	1.22	***	12,664	7.43	***	4062	2.30	***
	Std			15,214		531			2009			808		

Panel B: average daily number of trades

		Total		SIX		BS			CHI			TQ		
		Trades	FI	Trades	Ftaction (%)	Trades	Fraction (%)	Sig.	Trades	Fraction (%)	Sig.	Trades	Fraction (%)	Sig.
Pooled sample	Mean	128,644	1.85	91,530	72.34	6395	4.40	***	21,960	16.29	***	8759	6.98	***
	Std			2992		492			1147			401		
Stocks L	Mean	83,814	1.97	57,174	69.22	4632	4.96	***	15,692	18.04	***	6317	7.79	***
	Std			3539		727			1517			495		
Stocks M	Mean	29,393	1.76	21,532	74.98	1376	3.97	***	4659	14.73	***	1826	6.31	***
	Std			1509		265			642			201		
Stocks S	Mean	15,437	1.43	12,825	84.11	387	2.25	***	1610	9.89	***	616	3.75	***
	Std			1003		92			251			101		

The table shows market share figures in terms of trading volume (turnover in CHF) and number of trades for the Swiss exchange (SIX), BATS Europe (BS), Chi-X (CHI), Turquoise (TQ), and all trading venues (Total) over the sample period November 3, 2008, until June 30, 2010. Panel A shows the fraction of the different trading venues in the average daily trading volume. Panel B shows the market share in terms of the average daily number of trades over the sample period. Statistical significance for the mean differences of the market share between the Swiss exchange and the three MTFs is tested with a standard *t*-test. Source: Kohler and von Wyss [8]

$^{***}/^{**}/^{*}$Denotes significance at the 1 %/5 %/10 % level the number of trades is higher for all MTFs than the market share in terms of trading volume, it follows that the average trade size is lower on the MTFs than on the Swiss exchange.

An increase in the fragmentation index FI is related to an increase in dispersion on different venues. The results for the subsamples show that fragmentation increases for the higher capitalized stocks. The fragmentation index is 1.55 for the trading volume and 1.97 for the number of trades for Stocks L. Stocks M exhibit a lower degree of fragmentation for trading volume (FI = 1.44) and for the number of trades (FI = 1.76) and the highest concentration in trading can be found for Stocks S with FI = 1.26 for trading volume and FI = 1.43 for the number of trades. Figure 23.1 shows the development of the fragmentation index (FI) over the sample period together with the corresponding trading volume and the corresponding number of trades.

Figure 23.1 shows a steady increase in fragmentation over the sample period. For trading volume the fragmentation is rather stable until June 2009 and increases between June 2009 and June 2010. A temporary decrease of FI in December 2009, which is more pronounced for the trading volume than for the number of trades, coincides with a decrease in the overall trading activity reflected in total trading volume and total number of trades. Overall, the fragmentation is increasing for all subsamples; however, the increase is more pronounced for the higher capitalized stocks.

To assess market quality we calculate four liquidity measures that capture different dimensions of liquidity, namely relative spread, relative effective spread, dollar depth, and turnover.

Average liquidity measures across trading venues and subsamples are given in Table 23.3.

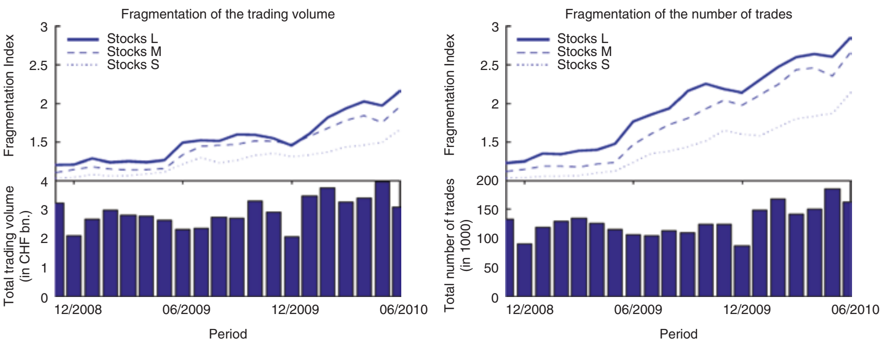


Fig. 23.1 Fragmentation of trading volume and number of trades. The figure shows the fragmentation index (FI) over the sample period November 3, 2008, until June 30, 2010, for total trading volume (*left panel*) and for the total number of trades (*right panel*). Additionally, the total trading volume (in CHF bn.) is shown in the *left panel*. The total trading volume is the monthly average of the aggregated daily trading volume on all trading venues. In the *right panel*, the total number of trades (in 1000) is shown, calculated as monthly average of the aggregated number of trades per day on all four trading venues. Source: Kohler and von Wyss [8, 9]

Table 23.3 Liquidity measures

RS		SIX (%)	BS (%)	Sig.	CHI (%)	Sig.	TQ (%)	Sig.
Pooled sample	Mean	0.16	0.34	***	0.38	***	0.33	***
	Std	0.09	0.46		1.07		1.08	
Stocks L	Mean	0.10	0.29	***	0.18	***	0.19	***
	Std	0.04	0.47		0.21		0.41	
Stocks M	Mean	0.15	0.35	***	0.29	***	0.33	***
	Std	0.05	0.39		0.86		1.34	
Stocks S	Mean	0.23	0.39	***	0.83	***	0.56	***
	Std	0.12	0.52		1.81		1.34	

RS^{eff}		SIX (%)	BS (%)	Sig.	CHI (%)	Sig.	TQ (%)	Sig.
Pooled sample	Mean	0.06	0.09	***	0.08	***	0.08	***
	Std	0.04	0.15		0.12		0.11	
Stocks L	Mean	0.04	0.07	***	0.05	***	0.05	***
	Std	0.01	0.14		0.06		0.09	
Stocks M	Mean	0.05	0.10	***	0.08	***	0.08	***
	Std	0.02	0.16		0.09		0.09	
Stocks S	Mean	0.08	0.12	***	0.14	***	0.13	***
	Std	0.05	0.14		0.19		0.15	

V		SIX	BS	Sig.	CHI	Sig.	TQ	Sig.
Pooled sample	Mean	9,767,894	358,340	***	1,464,092	***	589,678	***
	Std	13,782,305	1,031,524		2,736,155		1,043,423	
Stocks L	Mean	22,381,217	887,624	***	3,523,297	***	1,451,030	***
	Std	17,030,503	1,608,674		3,813,634		1,391,390	
Stocks M	Mean	4,203,483	125,381	***	569,065	***	209,806	***
	Std	3,534,994	249,377		752,690		262,516	
Stocks S	Mean	1,935,770	29,091	***	170,560	***	54,701	***
	Std	1,843,876	64,301		243,436		97,949	

D$		SIX	BS	Sig.	CHI	Sig.	TQ	Sig.
Pooled sample	Mean	182,093	28,128	***	52,316	***	32,204	***
	Std	393,843	78,010		121,840		52,583	
Stocks L	Mean	389,160	54,674	***	111,850	***	66,574	***
	Std	616,209	110,127		191,748		76,366	
Stocks M	Mean	92,245	11,341	***	27,021	***	18,443	***
	Std	53,250	15,476		23,172		15,719	
Stocks S	Mean	51,849	17,286	***	14,273	***	9305	***
	Std	34,308	68,220		16,986		10,374	

The table shows liquidity measures for the Swiss exchange (SIX), BATS Europe (BS), Chi-X (CHI), and Turquoise (TQ) across the subsamples. RS denotes the relative spread and RS_{eff} denotes the relative effective spread. V denotes the average turnover per hour in CHF and D$ is the dollar depth, measured as average posted volume on the bid and ask side of the order book. Additionally, statistical significance for the mean differences between the Swiss exchange and the three MTFs is tested with a standard *t*-test. Source: Kohler and von Wyss [8]

***/**/*Denotes significance at the 1 %/5 %/10 % level

The Swiss exchange provides the highest liquidity according to all liquidity measures. The relative spread for the pooled sample on Chi-X, BATS Europe, and Turquoise is 0.33–0.38 % which equals roughly two times the relative spread on the Swiss exchange. The relative effective spread (RS_{eff}) on the Swiss exchange for the pooled sample is 0.06 % which equals approximately one-third of the relative spread. The same proportion can be seen in RS_{eff} for Chi-X, BATS Europe, and Turquoise with RS_{eff} between 0.08 and 0.09 %. The fact that the relative effective spread is smaller than half the relative spread shows that trades are executed within the quote. The spread measures decrease with increasing market capitalization. Turnover V and dollar depth $D\$$ show a similar pattern, as they are higher on the Swiss exchange than on the MTFs and increase with market capitalization. The differences between the liquidity measures on the Swiss exchange and the MTFs are all highly significant.

According to the analyzed liquidity measures, Chi-X is the MTF with the highest market quality, followed by Turquoise and BATS Europe. Figure 23.2 shows how the liquidity measures evolve over the sample period. The upper panel shows the relative spread RS and the relative effective spread RS_{eff}, weighted with the corresponding turnover per trading venue. The lower panel shows turnover V and dollar depth $D\$$, both in log scales.

According to Fig. 23.2 the spread measures are decreasing over the sample period for all subsamples and the spreads for the higher capitalized stocks are consistently lower than for the smaller stocks. Turnover does not show a clear trend while dollar depth, especially for Stocks L, is increasing over time which is consistent with Foucault and Menkveld [2], who also find a deeper consolidated order book after the entrance of the MTF EuroSETS in the Dutch stock market. Figure 23.2 clearly shows an increase in market quality which coincides with a steady increase in fragmentation.

23.4 Information Processing

A central aspect in the analysis of fragmented markets is information processing, i.e., how information is incorporated into prices and which trading venue is leading. Two studies that analyze this question in the fragmented European equity market after the implementation of MiFID are Storkenmaier et al. [3] and Riordan et al. [4]. Storkenmaier et al. [3] analyze stocks that are traded on the LSE and Chi-X and find for the quote-based price discovery higher information shares for Chi-X (58.19 %), than for LSE (41.81 %), although LSE provides more liquidity. Furthermore, they analyze market reactions of LSE and Chi-X to Thomson Reuters newswire messages and find a shift of information processing towards LSE on days where positive news outweigh. Riordan et al. [4] also report quote-based information shares for Chi-X, which are higher (56.77 %) than for LSE (27.63 %) or other MTFs, like BATS (11.66 %) or Turquoise (3.94 %).

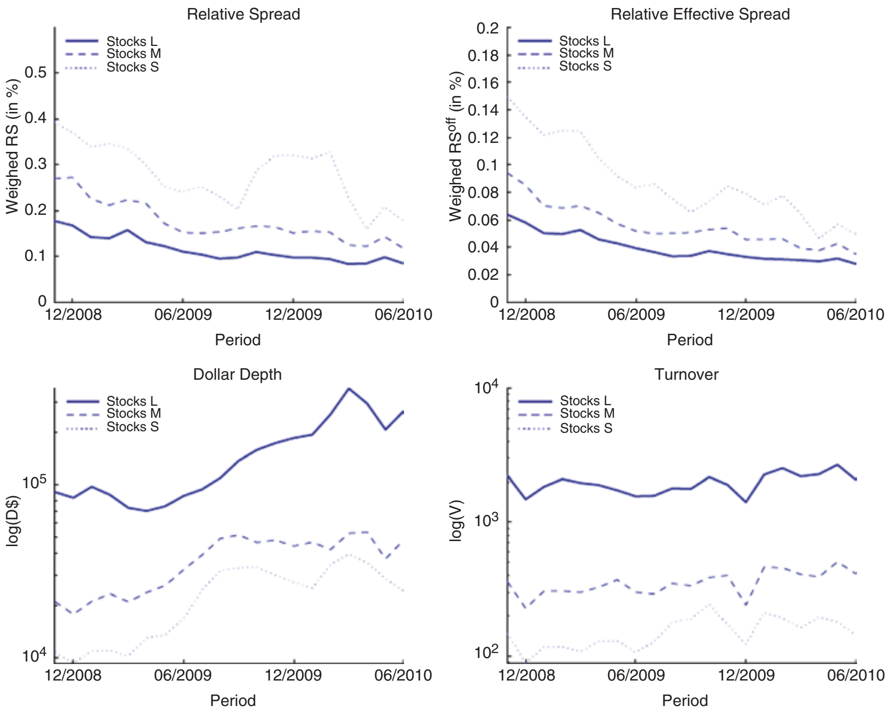


Fig. 23.2 Liquidity measures. The figure shows the development of the liquidity measures relative spread RS, relative effective spread RS_{eff}, dollar depth $D\$$, and turnover V over the sample period for all trading venues. RS and RS_{eff} are weighted with the corresponding turnover. $D\$$ and V are reported in log scale. Source: Kohler and von Wyss [8, 9]

Both studies apply Hasbrouck information shares (see Hasbrouck [5]) for the attribution of information shares to the different trading venues. Although information shares according to Hasbrouck is a widely used concept, there are two main drawbacks: First, information shares require equidistant data and, therefore, do not take the asynchronous nature of intraday data (e.g., order arrivals or order book changes) into account. Second, if there is contemporaneous correlation in the price innovations across different trading venues, the Hasbrouck information share of a market is not uniquely determined, but given in terms of upper and lower bounds. Typically, these bounds cover a wide range, which makes the clear identification of a leading venue impossible.

We also apply Hasbrouck information shares, but extend the analysis by an autoregressive conditional intensity (ACI) model according to Russell [6] as a new measure. We analyze information processing on the Swiss exchange and on Chi-X, which is the largest MTF competing with the Swiss exchange. By modelling the conditional intensities of the order arrivals, we can exploit the duration structure of the effective order arrivals without the loss of information that results from time aggregation. Therefore, we can incorporate typical characteristics of asynchronous order arrivals and we get unbiased point estimates for the information shares of the two trading venues, rather than just upper and lower bounds.

23.4.1 Hasbrouck Information Shares

Table 23.4 shows the average daily HIS per stock for the pooled sample and for two subsamples of large stocks (Stocks L) and small stocks (Stocks S).

The mean information share of the Swiss exchange for the pooled sample equals 53.25 %, which would indicate that the Swiss exchange has a higher information share than Chi-X. However, the median information share of the Swiss exchange is 48.16 %, which is slightly below 50 %. The problem of clearly identifying the leading venue in terms of the information share arises with the consideration of the upper and lower bounds of HIS. Figure 23.3 shows the estimated HIS together with the upper and lower bounds HIS_{up} and HIS_{low}, respectively.

HIS^{SWX} and HIS^{SWX} are calculated as mean of the respective upper and lower bounds HIS_{up} and HIS_{low}. This means that for the pooled sample the information share of the Swiss exchange lies between 40.47 % ($HIS_{low}{}^{SWX}$) and 66.04 % ($HIS_{up}{}^{SWX}$) and the information share of Chi-X between 33.96 % ($HIS_{low}{}^{CHI}$) and

Table 23.4 Hasbrouck information shares

	HIS^{SWX} (%)	HIS^{CHI} (%)	$HIS_{up}{}^{SWX}$ (%)	$HIS_{up}{}^{CHI}$ (%)	$HIS_{low}{}^{SWX}$ (%)	$HIS_{low}{}^{CHI}$ (%)
Panel A: pooled sample						
Mean	53.25	46.75	66.04	59.53	40.47	33.96
Median	48.16	51.84	63.94	67.55	32.45	36.06
Q75	67.62	63.50	77.49	81.23	58.80	45.87
Q25	36.50	32.38	54.13	41.20	18.77	22.51
Panel B: stocks L						
Mean	44.12	55.88	60.33	72.08	27.92	39.67
Median	41.34	58.66	58.87	75.92	24.08	41.13
Q75	51.65	67.23	68.21	85.44	35.74	49.10
Q25	32.77	48.35	50.90	64.26	14.56	31.79
Panel C: stocks S						
Mean	62.38	37.62	71.74	46.98	53.02	28.26
Median	60.92	39.08	71.96	49.07	50.93	28.04
Q75	81.86	55.73	86.46	71.10	78.18	41.15
Q25	44.27	18.14	58.85	21.82	28.90	13.54

The table shows the average daily mean, median, first, and third quartile of the Hasbrouck information shares (HIS) together with the *upper* and *lower* bounds (HIS_{up} and HIS_{low}) for the Swiss exchange (SWX) and for Chi-X (CHI) over the sample period January 1 to March 31, 2010. Panel A covers the pooled sample and Panel B and Panel C the subsamples Stocks L and Stocks S, respectively. The information shares are calculated as daily means of the *upper* and *lower* bound. Source: Kohler and von Wyss [9]

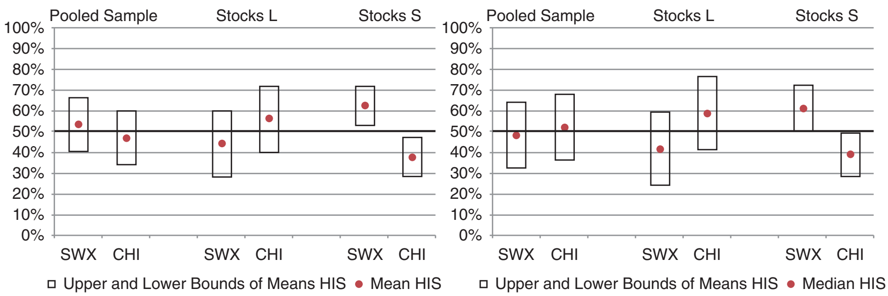


Fig. 23.3 *Upper* and *lower* bounds of Hasbrouck information shares. The figure shows estimated mean (*left panel*) and median (*right panel*) Hasbrouck information shares (HIS) for the pooled sample and the two subsamples Stocks L and Stocks S over the sample period January 1 to March 31, 2010. The range between the *upper* and *lower* bounds for the Swiss exchange (SWX) and Chi-X (CHI) is presented as rectangle. The arithmetic mean of the respective *upper* and *lower* bound is denoted by a *red dot*. Source: Kohler and von Wyss [9]

59.53 % (HIS_{up}^{CHI}), respectively. No trading venue has an information share which lies clearly above or below 50 %, which makes the identification of the leading trading venue for the pooled sample impossible.

The same holds true for subsample Stocks L. Although the mean and the median information share of Chi-X are larger than 50 %, according to the upper and lower bounds, a clear identification of the leading venue is not possible as the mean of HIS_{low}^{SWX} lies with 27.92 % below 50 % and the mean of HIS_{up}^{SWX} with 60.33 % above 50 %.

For subsample S the mean and median information share is higher for the Swiss exchange than for Chi-X with a mean information share of 62.38 % and a median information share of 60.92 % for the Swiss exchange. For this subsample the range between upper and lower bounds of HIS is disjoint, which allows the identification of the Swiss exchange as trading venue "who moves first."

Overall, the question which trading venue is actually leading in terms of Hasbrouck information shares cannot be answered conclusively. For the large caps some evidence is found that Chi-X is the leading market, which would confirm the results of Storkenmaier and Wagener [7] and Riordan et al. [4]. However, upper and lower bounds of HIS do not allow a clear identification of the leading venue. For the small caps evidence suggests that the Swiss exchange is the leading market.

23.4.2 Autoregressive Conditional Intensity Model

Based on the estimation of an ACI(1,1) model we calculate intensity-based information shares for the pooled sample and the two subsamples. Table 23.5 gives the results.

Table 23.5 Intensity-based information shares

	IIS^{SWX} (%)	IIS^{CHI} (%)
Panel A: pooled sample		
Mean	36.6	63.4
Median	33.6	66.4
Q75	52.8	83.7
Q25	16.3	47.2
Lead 95 %	7.1	46.4
Lead 99 %	3.6	42.9
Panel B: stocks L		
Mean	37.6	62.4
Median	42.1	57.9
Q75	57.3	84.0
Q25	16.0	42.7
Lead 95 %	14.3	57.1
Lead 99 %	7.1	57.1
Panel C: stocks S		
Mean	35.6	64.4
Median	31.1	68.9
Q75	44.6	83.3
Q25	16.7	55.4
Lead 95 %	0.0	35.7
Lead 99 %	0.0	28.6

The table shows intensity-based information shares estimated from the bivariate autoregressive conditional intensity (ACI) model for the intensity of order book changes of the Swiss exchange (SWX) and Chi-X (CHI). Panel A covers stocks from the pooled sample and Panel B and Panel C stocks from the subsamples Stocks L and Stocks S, respectively. Lead 95 % and Lead 99 % denote the fraction of stocks in the respective subsamples, where the intensity-based information share of one market is significantly higher than 50 % with a confidence level of 95 % and 99 %, respectively. Source: Kohler and von Wyss [9]

The intensity-based information share for Chi-X equals 63.4 % in terms of the mean and 66.4 % in terms of the median which means that for the pooled sample Chi-X is the leading market in terms of the intensity-based information share. The lead of Chi-X is highly significant for 42.9 % of the pooled sample, whereas the lead of the Swiss exchange is only significant for 3.6 % of the stocks. These findings are supported by the analysis of the two subsamples. For Stocks L the mean of IIS^{CHI} equals 62.4 % and for 57.1 % of the stocks in subsample Stocks L the lead of Chi-X is highly significant. The same holds true for subsample Stocks S with a mean IIS^{CHI}

of 64.4 %. However, the lead of Chi-X is only significant at the 1 % level for 28.6 % of the stocks. There is no stock in subsample Stocks S for which the Swiss exchange is significantly leading at the 1 or 5 % level.

Overall, we find strong evidence that Chi-X is the leading market in terms of the intensity-based information shares. Although the first quartiles of IIS^{CHI} lie below 50 % for the pooled sample and subsample Stocks L, the mean estimates, which in case of the intensity-based information shares are point estimates for the true values, lie well above the 50 % threshold and are confirmed by respective significance tests.

The findings from the analysis of the intensity-based information shares confirm our findings from the Hasbrouck information shares for subsample Stocks L, which suggested that Chi-X is the leading trading venue. The intensity-based information shares also confirm the lead of Chi-X for the second subsample Stocks S, where Hasbrouck information shares suggest a lead of the Swiss exchange.

23.5 Conclusion

The implementation of MiFID served as a catalyst for the emergence of MTFs in Europe which led to an increased fragmentation of liquidity in European equity trading. We investigate a sample of stocks that are listed on the Swiss exchange and the three MTFs Chi-X, BATS Europe, and Turquoise over a long-term sample that covers 20 months. We find no evidence for a deterioration of market quality in the aftermath of the implementation of MiFID. In contrast, there are significantly positive effects of the fragmentation on spread and depth measures, which are confirmed by the analysis of different subsamples.

Our study provides evidence that the fragmentation of trading in Swiss equity markets did not deteriorate market quality.

Previous studies have analyzed information processing after MiFID with the well-known Hasbrouck information shares. We also apply Hasbrouck information shares with inconclusive results. Evidence suggests that Chi-X is the leading trading venue for larger stocks, whereas for smaller stocks the Swiss exchange is still leading. However, overall the clear identification of the leading venue according to Hasbrouck information shares is not possible.

By applying an autoregressive conditional intensity model, we calculate intensity-based information shares, which take the effective irregular duration structure of order book changes into account. Furthermore, the autoregressive intensity model allows calculating statistically meaningful point estimates for the information shares of the respective trading venues. We find significant cross effects between the intensity processes of the Swiss exchange and Chi-X. Additionally, we provide evidence that Chi-X is the leading market in terms of intensity-based information processing irrespective of the market capitalization of the stocks.

References

1. EC (2004). Council Directive 2004/39/EC of 21 April 2004 on markets in financial instruments.
2. Foucault, T. and Menkveld, A. J. (2008). Competition for Order Flow and Smart Order Routing Systems. The Journal of Finance, 63(1):119–158.
3. Storkenmaier, A., Wagener, M., and Weinhardt, C. (2012). Public Information in Fragmented Markets. Financial Markets and Portfolio Management, 26(2):179–215.
4. Riordan, R., Storkenmaier, A., and Wagener, M. (2011). Do Multilateral Trading Facilities Contribute to Market Quality? Working Paper.
5. Hasbrouck, J. (1995). One Security, Many Markets: Determining the Contributions to Price Discovery. The Journal of Finance, 50(4):1175–1199.
6. Russell, J. R. (1999). Econometric Modelling of Multivariate Irregularly-Spaced High-Frequency Data. Working Paper.
7. Storkenmaier, A. and Wagener, M. (2011). Do we need a European 'National Market System'? Competition, Arbitrage, and Suboptimal Executions. Working Paper.
8. Kohler, A. and von Wyss, R. (2012). Fragmentation in European Equity Markets and Market Quality—Evidence from the Analysis of Trade-Throughs. Working paper.
9. Kohler, A. and von Wyss, R. (2012b). Where does Information Processing in a Fragmented Market Take Place?—Evidence from the Swiss Stock Market after MiFID. Working paper.
10. EC (2008). Council Directive 2008/10/EC of 11 March 2008 on markets in financial instruments.

Chapter 24
Takeover Regulation as Part of a Functioning Equity Market

Christian Zschocke

24.1 The Principles

Takeover law provides rules applicable to an offeror that aims at acquiring the shares from the shareholder of a listed target company by way of a public tender offer.

Takeover regulations are typically based on certain main guiding principles[1] for such offer process: the equal treatment of all holders of the same class of shares of the target company; the transparency of the offer proceeding and the obligation to provide the shareholders with comprehensive information required to enable them to reach a properly informed decision; as well as an adequate offer period so that the shareholders can decide without time pressure whether to accept the offer.

These protective rules aim at preventing certain takeover strategies that are considered to unfairly benefitting the offeror to the disadvantage of the other shareholders of the target company, e.g., front-end strategies rewarding shareholders that accept quickly an offer,[2] or insider trading.

---

[1] These principles can be found, *e.g.*, as General Principles in the European Takeover Directive or the City Code on Takeovers and Mergers in the United Kingdom ("City Code").

[2] The report on takeover and other bids by Professor Robert R. Pennington, EU Commission Working Document XI/56/74-E ("Pennington Report"), stated in point 88, p. 75, that the reason for requiring the offeror to offer all those persons to whom an offer is addressed exactly the same terms for their holdings is to prevent "*special arrangements for the benefit of holders of large blocks of securities by which they are paid rateably more for their securities than smaller investors*." The Legislative Materials for the German WpÜG, BT-Drs. 14/7034, p. 35, state (with respect to the rule stipulating that holders of the same class of securities of a target company shall be treated equally) that offers where the amount of consideration offered is staggered in accordance with the time of the declaration of acceptance, in order to cause a "*greyhound race*" of the holders of the securities, violate the general principle of equal treatment.

C. Zschocke (✉)
Morgan, Lewis and Bockius LLP, Bockenheimer Landstraße 4,
60323 Frankfurt am Main, Germany
e-mail: christian.zschocke@morganlewis.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_24

24.2 The Jurisdictions

24.2.1 The European Takeover Directive

The creation of a European takeover regulation took almost 30 years of legal argumentation[3] until it finally resulted in 2004 in the adoption of the European Takeover Directive ("Takeover Directive").[4]

The Directive provides for a general framework aiming at a takeover level playing field among member state legislations. It contains certain guidelines applying to mandatory and voluntary takeover offers for listed target companies. The objective is to enhance legal fairness and certainty for a takeover offer, in particular by ensuring a protection of the interests of the minority shareholders.[5]

The Takeover Directive provides for a mandatory offer once the offeror holds so much voting rights in the target company giving him or her control of the target company. It lays out requirements for the equitable price to be offered in the case of a mandatory offer. Member states shall ensure, following an offer made to all shareholders of the target company, if certain participation thresholds are fulfilled, (1) the right of the offeror to a squeeze-out of the remaining shareholders,[6] as well as (2) the right of the shareholders to a sell-out.[7] Member states are granted the option ("opt-in/opt-out") to incorporate either the stricter European rules on board neutrality[8] and/or the breakthrough rule[9] into national mandatory statutory law or to maintain a more lenient regime.[10]

---

[3] *Cf.* the Pennington Report (*ibid.*; footnote 2); a first legislative initiative failed when the European Parliament rejected a joint draft of Commission and Council on July 4, 2001.

[4] Directive 2004/25/EC of the European Parliament and of the Council of April 21, 2004, on takeover bids, Official Journal of the European Union, April 30, 2004, No. L 142/12.

[5] European Commission, Report to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the application of Directive 2004/25/EC on takeover bids (COM(2012) 347 final, dated June 28, 2012 ("2012 Report"), p. 2, paragraph 3.

[6] If the prerequisites for a squeeze-out are fulfilled, member states shall ensure that the offeror is able to require all the holders of the remaining securities to sell those securities at a fair price to the offeror.

[7] If the prerequisites for a sell-out are fulfilled, member states shall ensure that a holder of remaining securities of the target company is able to require the offeror to buy those remaining securities at a fair price under the same circumstances as provided for with respect to a squeeze-out.

[8] The board neutrality rule (Article 9 of the Takeover Directive) provides that during the bid period the board of the target company must obtain prior authorization from the general meeting of shareholder before taking any action which might result in the frustration of the offer.

[9] The breakthrough rule (Article 11 of the Directive) neutralizes pre-bid defences during a takeover by making certain restrictions (*e.g.*, share transfer or voting restrictions) inoperable during the takeover period and allows a successful offeror to remove the incumbent board of the target company and modify its articles of association.

[10] The Commission stated in its 2012 Report (*ibid.*; footnote 5), p. 3, paragraph 7) that 19 member states have transposed the board neutrality rule while 3 member states have transposed the breakthrough rule. About half of the member states allow companies who are subject to the board neutrality rule and/or breakthrough rule (by law or based on the articles of association of the company)

Member states had to implement the Takeover Directive until May 20, 2006. In its 2012 report on the application of the Takeover Directive ("2012 Report")[11] the Commission concluded that the Takeover Directive is working satisfactorily[12] and that similar rules already existed or were in the making at the national level prior to the adoption of the Directive.[13] A comparison with third countries[14] showed that takeover offer legislation in those countries is based on similar principles to those in the Takeover Directive, including a mandatory offer rule and a board neutrality rule (except for the USA). None of the investigated third countries has an equivalent to the breakthrough rule and only a limited number of these countries have rules regarding squeeze-out and sell-out rights.[15]

24.2.2 The Model of the UK

The City Code on Takeovers and Mergers ("City Code") had been developed since 1968 as a voluntary instrument of self-regulation and has become to a certain extent the model regulation of modern European takeover rules.

Following the implementation of the Takeover Directive,[16] the City Code received its own statutory basis in the UK.[17] The Companies Act 2006 provides the requirements under which a company may, by means of an "opting-in resolution," submit to the breakthrough rule of the Takeover Directive,[18] as well as the rules for the right

---

not to apply the rule when they are confronted with a takeover offer by an offeror who is not subject to the same rule (reciprocity).

[11] 2012 Report (*cf*. footnote 5).

[12] *Ibid*.; footnote 5, p. 9, paragraph 21.

[13] *Ibid*.; footnote 5, p. 3, paragraph 6. An external study conducted on behalf of the Commission included Austria, Belgium, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, the Netherlands, Poland, Portugal, Romania, the Slovak Republic, Spain, Sweden, and the UK.

[14] *Ibid*.; footnote 5, p. 5, paragraph 13, footnote 15: An external study conducted on behalf of the Commission (http://ec.europa.eu/internal_market/company/docs/takeoverbids/study/study_en.pdf) included the following third countries: Australia, Canada, China, Hong Kong, India, Japan, Russia, Switzerland, and the USA.

[15] *Ibid*.; footnote 14. The external study (which included the third countries Australia, Canada, China, Hong Kong, India, Japan, Russia, Switzerland, and the USA) concluded on p. 237: Three out of nine Major Non-EU Jurisdictions (China, Japan, and the USA) do not provide for an option to squeeze out minority shareholders following a successful takeover bid. Such jurisdictions may provide for alternative mechanisms permitting the exclusion of minority shareholders. Five out of nine Major Non-EU Jurisdictions (India, Canada, Japan, Switzerland, and the USA) do not provide for a minority shareholder right to force the majority shareholder to sell out their shares.

[16] According to Part 28, Section 943 (1), of the Companies Act 2006, the Takeover Panel must adopt rules giving effect to Articles 3.1, 4.2, 5, 6.1–6.3, 7–9, and 13 of the Takeover Directive.

[17] The City Code also has received a statutory basis in relation to the Isle of Man, Jersey and Guernsey.

[18] Sections 966 *et seq*. Companies Act 2006.

of the offeror to "squeeze out"[19] and the right of shareholders to "sell out"[20] following a takeover offer.

The City Code[21] contains rules for mandatory and voluntary offers. As a general rule, a mandatory offer must be made[22] to the holders of any class of equity share capital whether voting or nonvoting[23] when (1) any person acquires an interest in shares which (together with shares of persons acting in concert with it) carry 30 % or more of the voting rights of a company,[24] or (2) any person, together with persons acting in concert with it, is interested in shares which in the aggregate carry not less than 30 % of the voting rights but does not hold shares carrying more than 50 % of such voting rights and such person, or any person acting in concert with it, increases the percentage of its voting shares. The mandatory offer price must be, in respect of each class of share capital involved, in cash or be accompanied by a cash alternative at not less than the highest price paid by the offeror or any person acting in concert with it during the 12 months prior to the announcement of that offer.

The City Code requires the management of a target company to seek shareholders' approval for any action that may result in the frustration of a bid.[25] A Takeover Panel is supervising the takeover procedure.

24.2.3 The German WpÜG

The takeover regulation in Germany began with voluntary instruments of self-regulation that were eventually replaced by statutory law when the Securities Acquisition and Takeover Act ("WpÜG") became effective on January 1, 2002.

24.2.3.1 The 1979 Guidelines

The Exchange Expert Commission at the Federal Ministry of Finance (Börsensachverständigenkommission, "BSK") concluded in the 1970s that the regulation of public takeovers is necessary to support fairness and transparency in the

---

[19] Sections 974 *et seq*., 979 *et seq*. Companies Act 2006.

[20] Sections 974 *et seq*., 983 *et seq*. Companies Act 2006.

[21] Its general principles are the same as the general principles set out in the Takeover Directive.

[22] Rule 9.1 of the City Code provides, as an exception, that a mandatory offer will not be required where control of the target company is acquired as a result of a voluntary offer made in accordance with the City Code to all the holders of voting equity share capital and other transferable securities carrying voting rights.

[23] The offer must also be made to the holders of any other class of transferable securities carrying voting rights. Offers for different classes of equity share capital must be comparable.

[24] "Control" is defined as an interest, or interests, in shares carrying in aggregate 30 % or more of the voting rights of a target company, irrespective of whether such interest or interests give de facto control.

[25] Rule 21 of the City Code.

capital markets and to protect shareholders and investors. In January 1979, the BSK published voluntary takeover guidelines[26] ("BSK Guidelines").[27] The objective of the BSK Guidelines was to formulate basic rules ensuring in particular the equal treatment of the shareholders of a target company, the provision of sufficient information to these shareholders, and an orderly and fair execution of public tender offer proceedings. The BSK Guidelines never achieved practical relevance.[28]

24.2.3.2 The 1995 Takeover Code

On July 14, 1995, therefore, the BSK adopted the so-called Takeover Code (Übernahmekodex)[29] which replaced the BSK Guidelines.[30] Still an instrument of self-regulation, the Takeover Code had the purpose to ensure transparency of the takeover proceedings and a fair participation of the minority shareholders in the setting of takeover prices, ensuring a fair and transparent takeover proceeding based on full information.

The Takeover Code provided for a mandatory offer which gave all shareholders of the target company the opportunity of a fair exit by obligating the offeror to extend an offer to all shareholders if the control of the company had changed.[31] Despite some criticism,[32] the Takeover Code has received practical application in the German capital market, considering the large number of tender offers that complied with it.[33] Its primary shortcoming was the lack of sanctions and enforceability, due to its legal construction as a voluntary code of conduct. In February 1999, the BSK concluded therefore that the takeover rules required the

---

[26] Leitsätze für öffentliche freiwillige Kauf- und Umtauschangebote bzw. Aufforderungen zur Abgabe derartiger Angebote in amtlich notierten oder im geregelten Freiverkehr gehandelten Aktien bzw. Erwerbsrechten (Guidelines for Public Voluntary Purchase and Exchange Tender Offers, as well as Invitations to Issue Such Offers, for Shares or Rights Traded on the Official Market or the Regulated Free Market).

[27] Text in German language published in Fleischer/Kalss, Das neue Wertpapiererwerbs- und Übernahmegesetz, Munich, 2002, p. 197 *et seq*.

[28] Schuster/Zschocke, Übernahmerecht/Takeover Law, Frankfurt am Main, 1996 (with supplement March 1998), p. 48 *et seq*.

[29] Bilingual edition, with annotations: *ibid*.; footnote 28, p. 74 *et seq*.

[30] *Ibid*.; footnote 28, p. 53. On October 16, 1997, the BSK adopted modifications to the Takeover Code, mainly concerning the mandatory offer and the related pricing rules, which were published on November 28, 1997, and entered into effect on January 1, 1998.

[31] Until the modifications of November 28, 1997 (*cf*. footnote 30), the control threshold pursuant to the Takeover Code was originally fixed at 50 % of the voting rights of the target company. According to the revised provision, control could be obtained below the previous threshold. With respect to shareholder resolutions, for instance, control was presumed only if the potential offeror obtained at least a share of voting rights that would have constituted a percentage of voting rights equal to at least 75 % of the share capital present and entitled to vote at each of the three preceding shareholders' meetings of the target company.

[32] *Cf., e.g*., Kallmeyer, in: ZHR 161(1997), p. 435 *et seq*.

[33] Schuster, in: Zschocke/Schuster, Bad Homburger Handbuch zum Übernahmerecht, Heidelberg, 2003, Part A, point 27.

binding force of a statute.[34] The call for a takeover statute was influenced by the groundbreaking hostile takeover offer by which Vodafone Airtouch plc took over Mannesmann AG in 2000. Although both parties had acknowledged to adhere to the Takeover Code as basis for the takeover, neither party fully complied with its rules during the takeover procedure.[35]

24.2.3.3 The 2002 German Securities Acquisition and Takeover Act ("WpÜG")

The WpÜG became effective on January 1, 2002, and aims at providing market participants a binding, statutory framework in line with international standards.[36]

The WpÜG lays down rules for three different types of offers for the acquisition of securities that were issued by a German-listed target company[37]: (1) takeover offers, (2) mandatory offers, and (3) voluntary tender offers. The distinction between the three types of offers is based on the acquisition of "control," which is defined as the holding of at least 30 % of the outstanding voting rights of the target company.[38] A takeover offer aims at the acquisition of control; a mandatory offer is required[39] once the offeror has acquired control of the target company. Voluntary tender offers are offers that are neither takeover offers nor mandatory offers.

The offeror is generally free to determine the amount of the consideration. In the case of a takeover or mandatory offer,[40] the higher of the following minimum amounts is required: (1) the volume-weighted average stock price ("VWAP") of the target company during the 3-month period immediately preceding the announce-

[34] Legislative Materials, BT-Drs. 14/7034, p. 27. The WpÜG is designed to ensure a fair and orderly procedure, greater transparency during the public tender offer procedure giving shareholders and employees comprehensive information rights, as well as equal treatment for minority shareholders.

[35] Zschocke/Rahlf, in Wegerich, Business Laws of Germany, 2012 edition, Volume 1, Chapter 2:2, p. 145 *et seq*.

[36] Legislative Materials, BT-Drs. 14/7034, p. 28.

[37] The WpÜG primarily applies to German stock corporations and partnerships limited by shares whose securities are admitted to trading on a domestic organized market. Certain provisions of the WpÜG apply to (cross-border) European offers in cases where the shares in domestic or foreign target companies are quoted on an organized market outside the target company's country of residence. Details are set out in the WpÜG and the WpÜG Applicability Regulation (WpÜG-Anwendbarkeitsverordnung).

[38] For the analysis of whether a shareholder holds a controlling stake in a target company within the meaning of the WpÜG not only the shares legally owned by the shareholder count towards the 30 % threshold but also voting rights that are attributed to the shareholder.

[39] The WpÜG and the WpÜG Offer Regulation (WpÜG-Angebotsverordnung) set forth the requirements under which, apart from the possibility of disregarding certain voting rights, an acquirer of control can be released from the obligation to make a mandatory offer by way of an exemption decision rendered by the BaFin in its discretion after having considered the interests of the participants involved (the offeror and the shareholders of the target company). The BaFin may issue its exemption decision subject to conditions to address subsequent changes in the facts underlying the acquirer's application.

[40] Details are set forth in the WpÜG and the WpÜG Offer Regulation.

ment of the offer or acquisition of control, or (2) the price paid or agreed for any previous acquisitions of shares in the target company during the 6-month period prior to the publication of the offer document.

The minimum pricing rules are an outflow of the principle of equal treatment of shareholders. The management of the target company is bound by certain neutrality obligations[41] during the time period between the announcement of the offer or acquisition of control and the publication of the outcome of the offer in the acceptance period.[42] The WpÜG provides for the possibility to "squeeze out" minority shareholders and a "sell-out" right of minority shareholders, following a takeover or mandatory offer.

The regulatory authority supervising offers pursuant to the WpÜG is the BaFin.

433 public tender offers pursuant to the WpÜG were announced until the end of 2015, of which 411 offers were completed. These completed offers include 202 takeover offers, 144 mandatory offers and 65 voluntary tender offers (Fig. 24.1). The BaFin prohibited a small number of offers. Until the end of 2015, 22 offers were not completed, predominantly as a result of the failure to fulfill conditions stipulated by the offer or such as reaching a minimum acceptance threshold.

24.2.4 Switzerland

Takeover regulation has been introduced in Switzerland, a non-EU member state, in 1995.[43]

These regulations apply to public tender offers[44] for holdings in companies with registered office in Switzerland whose equity securities are listed in whole or in part

[41] Subject to certain limitations, the management board must abstain from any acts that could frustrate the success of the offer. The WpÜG lays out four exemptions to the neutrality obligations. The prohibition to frustrate the offer does not apply to (1) acts that a prudent and diligent manager of a company not affected by a takeover offer would also have taken; (2) the search for a competing offer; (3) acts taken with the consent of the supervisory board; and (4) defensive acts of the management board approved by the shareholder meeting in advance.

[42] The WpÜG provides that a target company may provide in the articles of association that the neutrality rule and/or breakthrough rule under the Takeover Directive shall apply ("opt-in"); to ensure a level playing field for cross-border takeovers, a reciprocity clause (pursuant to such reciprocity clauses, provisions in the articles limiting the effectiveness of share transfer restrictions or the ability of the company's management to take defensive measures against takeovers shall be subject to the reservation that the offeror in the instant case or a company controlling this offeror is subject to a regulation equivalent to the provisions of the WpÜG for the case of its own potential takeover) may be included in the target company's articles of association.

[43] After the amendment of the 1995 Federal Stock Exchange Act, Swiss takeover law is governed primarily by several regulations: (1) the Federal Financial Market Infrastructure Act ("FMIA"); (2) the Financial Market Infrastructure Ordinance ("FMIO"); (3) the Swiss Financial Supervisory Authority's Financial Market Infrastructure Ordinance ("FMIO-FINMA"); and (4) the Ordinance of the Swiss Takeover Board on Public Takeover Offers ("TOO").

[44] The FMIA contains a section concerning "public acquisition offers" but does not clearly define the term "public."

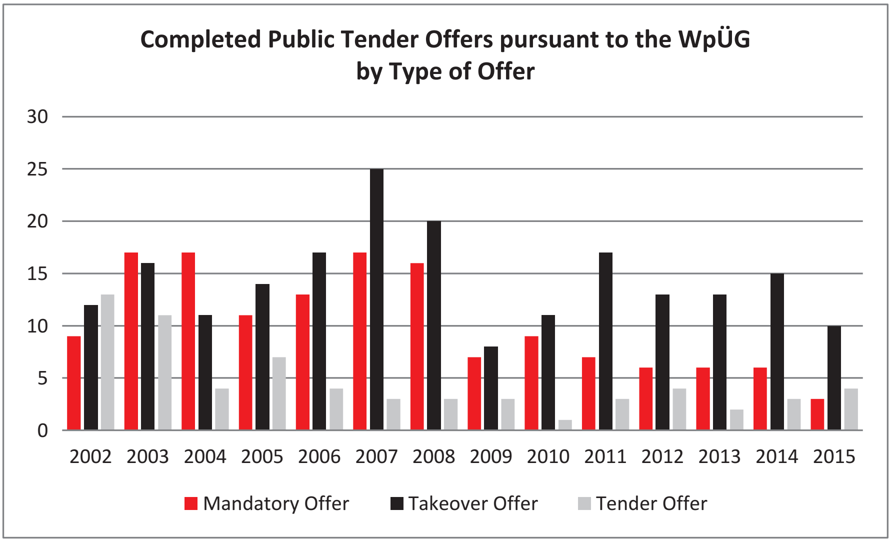


Fig. 24.1 Completed public tender offers pursuant to the WpÜG by type of offer

in Switzerland or in companies with registered office abroad whose equity securities are mainly listed[45] in whole or in part in Switzerland.[46]

Swiss takeover regulation distinguishes between mandatory and voluntary tender offers. A target company may at any time adopt a provision in its articles of association (so-called Opting-Out Clause) according to which an acquirer of control (*Übernehmer*) shall not be under the obligation to make a mandatory offer, provided that this does not prejudice the interests of shareholders.[47]

Whosoever, directly, indirectly or acting in concert with third parties, (i) acquires equity securities which, together with the equity securities already owned, exceed the threshold of 33 1/3 % of the voting rights[48] of a target company (target companies may increase the threshold in their articles of association to up to 49 % of the voting rights), whether or not such rights may be exercisable, or (ii) who, directly, indirectly or acting in concert with third parties, held, on February 1, 1997, equity securities which granted him or her control of more than 33 1/3 % but less than 50 %

[45] The FMIO provides that the equity securities of a company with registered office in a foreign country shall be deemed "mainly listed" in Switzerland if the company must fulfill at least the same duties for the listing and maintenance of the listing on a stock exchange in Switzerland as a company with registered office in Switzerland.

[46] The FMIA provides that if, in connection with a public acquisition offer, Swiss and foreign law are simultaneously applicable, then application of the provisions of Swiss law may be relinquished to the extent that the application of Swiss law would result in a conflict with the foreign law and the foreign law provides a protection of the investors that is equivalent to that provided under Swiss law.

[47] Art.125 (3), (4) FMIA.

[48] The threshold is to be calculated on the basis of the total number of voting rights according to the entry in the commercial register, Art. 34 (1) FMIO-FINMA.

of the voting rights of a target company, and who acquires equity securities and thereby exceeds the threshold of 50 % of the voting rights, must make a mandatory offer[49] for all listed equity securities of the target company.[50] The mandatory offer price must be at least as high as the higher of the following amounts[51]: (1) the stock exchange price[52] or (2) the highest price paid[53] by the acquirer for equity securities of the target company in the preceding 12 months.[54]

Public tender offers[55] are subject to supervision regulation by the Takeover Board ("TOB"). Once the offer has been published, the target company shall notify the TOB in advance about any defensive measure that it is considering.[56]

An offeror who, upon expiration of the offer period, holds more than 98 % of the voting rights of the target company may, within 3 months, petition the court to cancel the outstanding equity securities.[57]

24.2.5 The 2014 Sika/Saint-Gobain Case

The controversial nature of Swiss Opting-Out Clauses is highlighted by the recent case Sika/Saint-Gobain.[58]

The articles of association of Sika AG ("Sika") contain an Opting-Out Clause, according to which an acquirer of shares of Sika shall not be obliged to make a mandatory offer, but also a clause[59] providing for a limitation on the transferability of registered shares of Sika. Sika's shareholder Schenker-Winkler Holding AG

---

[49] The FMIA does not contain a definition of the term "control."

[50] The Takeover Board may grant exceptions from the mandatory offer obligation in certain justified cases; *cf.* Art. 136 (1), (2) FMIA, Art. 40 *et seq*. FMIO-FINMA.

[51] Art. 135 (2) FMIA.

[52] Art. 42 FMIO-FINMA provides, inter alia, that this is the volume-weighted average price of all on-exchange transactions executed during the 60 trading days prior to publication of the offer or the advance announcement, as the case may be, which must be adjusted to take into account any sizable fluctuations owing to special events.

[53] Art. 43 FMIO-FINMA.

[54] The so-called best price rule, as laid down in Art. 10 TOO, provides that if the offeror acquires equity securities of the target company in the period from the publication of the offer until 6 months after the additional acceptance period at a price that exceeds the offer price, it must offer this price to all recipients of the offer. The best price rule also applies to the acquisition of financial instruments and to offers relating to such instruments.

[55] Public tender offers which are not qualified as mandatory tender offers are considered as voluntary tender offers.

[56] Art. 35 TOO. Art. 36, 37 TOO lay out certain unlawful or inadmissible defensive measures.

[57] Details are provided in Art. 137 FMIA, Art. 120 *et seq*. FMIO.

[58] TOB decision 594/01 of March 5, 2015, and TOB decision 598/01 of April 1, 2015.

[59] According to that clause, the board of Sika can reject an acquirer of registered shares as shareholder to the extent that such acquirer's number of registered shares exceeds 5 % of the total number of registered shares registered with the commercial register.

("SWH") holds a total of 52.62 % of the voting rights and 16.43 % of the share capital.[60] SWH is wholly owned by the family Burkard.

Family Burkard sold its SWH shares in 2014 to the French group Compagnie de Saint-Gobain ("Saint-Gobain").[61] The TOB confirmed the validity of the Opting-Out Clause[62] and determined the applicability of the Opting-Out Clause to the proposed acquisition of SWH holding that, therefore, Saint-Gobain and persons acting in concert with Saint-Gobain are not obliged to submit a public offer for the listed shares of Sika.[63] While the TOB affirmed its competence to assess the validity and content of an Opting-Out Clause and the related issue whether an offer obligation pursuant to stock exchange law exists, the TOB denied its competence to independently assess, pursuant to corporate law, the validity and content of a clause providing for a limitation of the transferability of shares which assessment falls into the jurisdiction of the civil court.[64]

In the latter respect, the takeover created an additional controversy as Sika has denied SWH to vote more than 5 % during its shareholders' meetings since 2015, arguing that the selling shareholders in SWH and Saint-Gobain would be considered acting as a group.[65] This dispute is pending.[66]

24.2.6 The US Solution

Takeovers in the USA are governed by law at federal level and at state level, and are mainly based on securing full information to the shareholders in a takeover situation (disclosure approach). Of course, the extensive case law has also to be taken into account.

---

[60] The share capital of Sika comprises registered shares with restricted transferability, which are not listed anymore since 2003, and bearer shares which are listed in the main standard segment at the SIX Swiss Exchange.

[61] On December 22, 2014, Saint-Gobain assigned all rights and obligations of the SPA to a company controlled by it.

[62] TOB decision 594/01 of March 5, 2015, p. 8, 12.

[63] TOB decision 598/01 of April 1, 2015, p. 10, 13; the TOB decision was confirmed by the decision of the FINMA Takeover Committee of May 4, 2015 (p. 10 *et seq*., 17) and the judgment of the Swiss Federal Administrative Court of August 27, 2015 (B-3119/2015).

[64] *Ibid.*; footnote 63, p. 8.

[65] *Cf., e.g*., media releases of Sika dated January 26, 2015, April 15, 2015, July 24, 2015, or April 12, 2016.

[66] Sika stated in media releases dated April 7, 2015 and June 11, 2015 that in the proceedings concerning SWH's requests with regard to the restriction of the voting rights, as well as for an *ex parte* order prohibiting any restrictions of SWH's voting rights at the general meeting on April 14, 2015, SWH's requests were denied by the Cantonal Court of Zug and that SWH's appeal was denied by the Superior Court of Zug which held that the issue of the restriction of the voting rights shall be decided in ordinary proceedings. Sika stated in a media release dated October 2, 2015 that SWH has challenged certain decisions of the extraordinary shareholders' meeting of July 24, 2015, mentioning that the principal issue in both proceedings is whether the transfer restriction as set forth in Sika's articles applies to the intended sale of family Burkard's stake in Sika to Saint-Gobain.

At federal level the public trading of securities is regulated by the Securities Exchange Act of 1934 ("Exchange Act"), which applies to all listed companies. The primary regulator of the US securities markets is the Securities and Exchange Commission ("SEC"). Certain rules in the Exchange Act are aiming at the protection of shareholders of a target company whose shares are registered on a US securities exchange, and listed with the SEC, by means of mandatory full and fair disclosure of information regarding tender offers in filings.[67]

Such rules were in 1968 incorporated into the Exchange Act by the Williams Act,[68] in order to regulate tender offers and takeover bids. Under the Exchange Act, parties who will own more than 5 % of a class of the company's securities after making a tender offer[69] for (or who acquired more than 5 % of a voting class of) a company's securities registered under the Exchange Act are required to make certain filings of disclosure statements with the SEC within 10 days.

A mandatory tender offer is not required under US takeover law.[70] Antifraud provisions in the Exchange Act prohibit fraudulent, deceptive, and manipulative acts in connection with the tender offer.

In addition to takeover law at federal level, many states in the USA adopted their own provisions for takeovers, which in parts deviate from each other significantly. These provisions were adopted predominantly as reaction to hostile takeovers in the 1980s in order to prevent negative implications such as redundancies and tax shortfalls.[71]

24.3 Outlook

The ever-developing capital markets will lead to a continuous further evolvement of the takeover rules on the national or international level.[72] Takeover rules will remain an integral part of a functioning equity market and mature together with it.

[67] Schuster, in: Zschocke/Schuster (*ibid.,* footnote 37), Part A, points 72 *et seq*.

[68] The Williams Act ("An Act providing for full disclosure of corporate equity ownership of securities under the Securities Exchange Act of 1934") added a number of provisions to Sections 13 and 14 of the Exchange Act addressing beneficial ownership disclosure, tender offers, and changes in control, including Sections 13(d) and 13(e) [15 U.S.C. 78 m(d)–(e)]; and Sections 14(d) and 14(e) [15 U.S.C. 78n(d)–(e)] (Commission Guidance on Mini-Tender Offers and Limited Partnership Offers, Part I, in: footnote 2).

[69] The term "tender offer" has never been defined in any statutory provision or rule (*cf*., footnote 68, Part I).

[70] Schuster, in Zschocke/Schuster (*ibid*.; footnote 35), Part A, point 73.

[71] Schuster, in Zschocke/Schuster (*ibid*.; footnote 35), Part A, point 75.

[72] EU member states have the competence to introduce additional measures which go beyond the requirements of the Takeover Directive as long as the Directive's general objectives are observed.

24.3.1 *"Acting in Concert"*

This will apply for instance to the question of which persons are considered to be "acting in concert" for takeover purposes. As the holdings of voting rights of "persons acting in concert" with the offeror are attributed to the offeror and added to the offeror's holdings, the concept of "acting in concert" is decisive for calculating whether the control threshold has been crossed and, as a consequence, whether an obligation to make a mandatory offer has arisen. The European Securities and Markets Authority ("ESMA") has published a public statement regarding information on shareholder cooperation and acting in concert under the "Takeover Directive."[73] The statement includes a "White List" of activities, in order to provide legal certainty to institutional investors as to the extent to which they can cooperate with each other without being regarded as "acting in concert" and running the risk of having to make a mandatory offer.[74]

24.3.2 *Creeping In*

Another controversial issue is the use of the so-called creeping-in strategy. An offeror using such approach would acquire an initial stake in the target company close to the applicable mandatory offer threshold and then extend an offer with no or only an insignificant offer premium on the statutory minimum price. As a result, such offeror frequently can pass the control threshold without major investment and without purchasing all minority shares with the usual premium. Afterwards, the offeror is no longer obliged to make a mandatory offer when it further increases its share in the target company.

Some market participants hold the view that under such "creeping-in" strategy, the offeror would not give minority shareholders a fair chance to exit the company and share in the control premium, and that this technique is not in line with the objective of the Takeover Directive to protect minority shareholders in situations of change of control.[75] The European Commission points out at examples in national legislation, such as additional mandatory offer thresholds,[76] to prevent the use of this strategy.

---

[73] ESMA/2013/1642, dated November 12, 2013. According to Art. 2 (1) (d) of the Takeover Directive, "persons acting in concert" shall mean natural or legal persons who cooperate with the offeror or the target company on the basis of an agreement, either express or tacit, either oral or written, aimed either at acquiring control of the target company or at frustrating the successful outcome of a bid.

[74] *Ibid.*; footnote 73, p. 5 *et seq.*

[75] *Ibid.*; footnote 5, p. 10, paragraph 25.

[76] *Ibid.*; footnote 5, p. 10, footnote 38, pointing to an overview on p. 130 of the External Study. Also the ESMA public statement includes an overview on national control thresholds, alternative primary thresholds, additional mandatory bid thresholds, and secondary thresholds in the concerned member states of the EU, as well as in Iceland (*Ibid.*; footnote 73, Appendix B, p. 11 *et seq.*).

The Commission may take steps to discourage the use of this technique across the EU.[77] The European Parliament took the view that national competent authorities should discourage such creeping-in techniques.[78] A legislative initiative in Germany for the introduction of additional mandatory offer thresholds was rejected by Parliament.[79]

24.3.3 "Parallel" Acquisitions of Securities (the McKesson/Celesio Case)

So-called parallel acquisitions have also become a field of legal discussion. The obtaining of corporate control can be influenced by the acquisition of nonvoting securities issued by a third party, such as convertible bonds, which carry for their holder the possibility of acquiring voting shares in the target company.[80] If the acquisition of such instruments, and the price paid or agreed for such, is not considered in the determination of the statutory minimum offer price, this creates a possibility for an offeror to avoid compliance with the minimum pricing rules. In such case, minority shareholders are unable to share in the control premium. In view of the objectives of equal treatment of shareholders and ensuring that shareholders may share in a fair way in the value of the target company, takeover regulation should cover at least "parallel" acquisitions carried out in the direct context of a takeover offer.

24.3.4 System of Legal Protection

Any regulation is only complete if it is complemented by an adequate system of legal protection. This raises numerous questions, including whether the making of a mandatory offer should be enforced also by the competent supervisory authority

[77] *Ibid.*; footnote 5, p. 10, paragraph 25.

[78] Resolution of May 21, 2013 (2012/2262(INI)), paragraph 15.

[79] Legislative Proposal, BT-Drs. 17/3481 dated October 27, 2010.

[80] In the context of a takeover offer under the German WpÜG by an offeror (belonging to the McKesson group) for shares of the target company Celesio AG, the offeror bought convertible bonds in order to convert them in the context of the offer to pass the target 75 % threshold. The Frankfurt Regional Court held in a judgment dated December 2, 2014 (3-05 O 44/14) that, in the context of determining the minimum offer consideration, acquisition prices for convertible bonds are irrelevant for determining the prior acquisition price for new shares created by conversion of the convertible bonds - at least in the case of an offeror's "derivative acquisition" of already issued convertible bonds from a seller. That judgment was set aside by the Frankfurt Higher Regional Court with a remarkable judgment dated January 19, 2016 (5 U 2/15), focusing closely on the economic purpose behind the acquisition of the convertible bonds in the specific case. However, in view of the importance of the applicability of the provision of the WpÜG at issue, the Frankfurt Higher Regional Court allowed an appeal which is currently pending before the German Federal Supreme Court.

and/or by the shareholders if the offeror does not submit an offer document although the legal requirements for a mandatory offer are fulfilled. Current takeover regulation typically provides that, as long as the offeror does not make the mandatory offer, the target company's shareholders cannot exercise their right to exit from the target company by accepting the offer.

The German Federal Supreme Court recently held[81] that the minority shareholders of a German target company are entitled neither to a claim for a consideration for their shares nor to a claim for damages if an acquirer of control does not comply with the statutory obligation to submit an offer document and to make a mandatory offer.[82] This raises an issue before the background of the target of a functioning equity market and the principle of the protection of the interests of the shareholders of the target company. Consideration should be given, therefore, to the introduction of rules providing legal protection to shareholders and supervisory authorities in this respect. The development of a functioning and balanced system of legal protection is one of the most difficult legal challenges to be worked on going forward.

24.3.5 The Role of Employees in a Takeover Situation

Employees of the target company are not considered direct parties to a takeover but rather "stakeholders" affected by the transaction. The Takeover Directive foresees that employees are provided with information on a takeover offer and that they may also provide their views on such offer.[83] An offeror typically is obliged to disclose in the offer document its intentions with regard to the future of the target company and the jobs of its employees and management.[84] The management and the supervisory board of the target company shall prepare a reasoned statement on the offer, on the effects of the implementation of the offer including on employment, as well as on the offeror's strategic plans for the target company and their likely repercussions on employment.[85] They shall communicate that statement to the representatives of the target company's employees or, where there are no such representatives, to the employees themselves. While some commentators plead for more involvement of the employees (or their representatives) in a takeover proceeding, the need for trans-

---

81 Judgment dated June 11, 2013, II ZR 80/12.

82 The court found, among other things, that the WpÜG has the purpose of creating framework conditions and, therefore, is primarily capital market law oriented, and that the provision stipulating the obligation to submit an offer document and make an offer is not a statute that is intended to protect another person.

83 Art. 6 (1), (2), (3) (i), Art. 8 (2) and Art. 9 (5) of the Takeover Directive.

84 Art. 6 (3) (i) of the Takeover Directive.

85 Art. 9 (5) of the Takeover Directive.

action certainty and the fact that a takeover concerns changes on ownership level of the target company, however, speak for continuing with the current approach limited to employee information.

24.4 Conclusion

Takeover regulation has become a generally accepted and integral component of a functioning equity market. It will further evolve and mature, hand in hand with the development of the capital markets.

Chapter 25
Implementing MiFID2: The View of a Cash Equities Trading Venue

Miroslav Budimir

25.1 Introduction

It has been more than 10 years that the European financial market has seen one of its most important pieces of legislation: the Markets in Financial Instruments Directisve (MiFID). Most of its provisions have been implemented in 2007. Its major achievement was the breakup of national monopolies in share trading. MiFID allowed competition in European equities trading, resulting in a sharp increase in the number of trading platforms. Apart from the desired outcome to boost competition, MIFID had also some unforeseen consequences. These partly resulted from the fact that the legislative text was written before the introduction of various technology changes, such as algorithmic trading and high-frequency trading (HFT). There have also been unintended consequences in the form of the creativity of market participants to find and exploit loopholes, e.g. the ability to trade outside the regulated environment. These developments were overshadowed by the Financial Crisis of 2008. MiFID was reviewed in the first half of the 2010s. As a result, the MiFID review addresses the market structure issues, but due to the crisis also focusses on the strengthening of transparency, stability and integrity of financial markets.

Before exploring these issues in detail, some clarity about the used terminology might be useful. Often, the terms MiFID or MiFID2 are used both for the whole legislative initiative including all relevant legislative texts and only for a respective Directive (for which the letter "D" stands). Figure 25.1 illustrates this structure. The original MiFID (for our purposes dubbed "MiFID1") is displayed on the left-hand side. It consists of a framework Directive[1] that outlines the general regulatory

[1] EU Parliament and Council [1].

M. Budimir (✉)
Deutsche Börse AG, Mergenthalerallee 61, 65760 Eschborn, Germany
e-mail: miroslav.budimir@deutsche-boerse.com

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_25

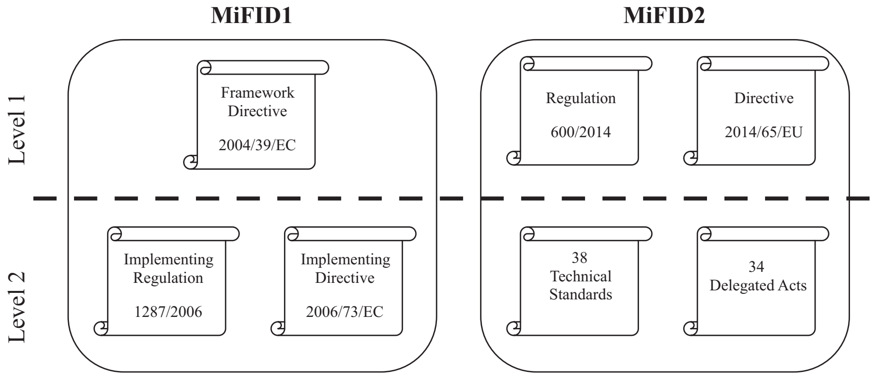


Fig. 25.1 The structure of MiFID1 and MiFID2 at a glance

approach ("Level 1") and two follow-up texts outlining further technical details ("Level 2"). The latter are subdivided into a Regulation[2] ("Implementing Regulation"), and another Directive[3] ("MiFID Implementing Directive", or Level-2 Directive). In this context it is useful to know that the difference between an EU "Regulation" and a "Directive" is that the former is directly valid throughout the Union, whereas the latter has to be implemented in the form of a national law in each respective member state. Throughout this article we will mean the left-hand part of Fig. 25.1 when referring to the original MiFID and dub this complex as "MiFID1", whereas the MiFID revision is "MiFID2". The latter comes in two flavours: A Level 1 Regulation[4] and a Level 1 Directive.[5] In this context it is noteworthy that "MiFID2" means more than just Directive 2014/65/EU, as it also consists of a multitude of Level 2 texts in the form of technical standards and delegated acts. In the following, when we mean specific texts, we will refer to their specific designations, such as 2014/65/EU.

The aim of this chapter is to shed some light on the development of MiFID and especially its review. We will focus predominantly on cash equity trading, meaning that other MiFID-related topics such as market data and access considerations, derivatives or commodities will not be tackled. In Sect. 25.2, we will highlight the impact of MiFID1 on the current market structure and some unforeseen and undesired consequences. In Sect. 25.3 we investigate the key changes brought by MiFID2 from the perspective of cash equity trading. In Sect. 25.4 we investigate the consequences of the new rules for implementation.

---

[2] EU Commission [2].

[3] EU Commission [3].

[4] EU Parliament and Council [4].

[5] EU Parliament and Council [5].

25.2 MiFID1: From the Concentration Rule to Electronic Trading

25.2.1 The Development of MiFID1

MiFID1 was developed in the first years of the 2000s and entered into force in 2004, but most of its provisions became effective on 1 November 2007. MiFID1 was preceded by the Investment Services Directive[6] of 1993, which has introduced the concept of remote membership, and thus paved the way for the introduction of the electronic stock exchange order books in Europe. In this step, the order execution process, that was until then handled manually by humans (e.g. Kursmakler or specialists), was automated. These developments can thus be seen as the ***first wave of automation of equities trading in Europe***. A ***second wave*** followed during the 2000s, when market participants transformed the manual order processing (order generation, -submission and -management.) into an automated process.

25.2.2 Competition: The Holy Grail for MiFID1

At the time that MiFID1 was developed, the focus of regulators was to create a single EU market for financial services.[7] One of the major goals of MiFID1 was the harmonisation of securities trading in Europe. Behind this background, investor protection and competition were to be strengthened.

One of the results of improved competition was the end to the concentration rules that existed throughout Europe in various forms.[8] On the one hand, MiFID1 introduced the concepts of “multilateral” and “bilateral” trading. The multilateral execution venues consisted of traditional exchanges (“regulated markets”) and of the newly created “multilateral trading facilities” (MTFs). Both types are labelled as “trading venues”. For bilateral trading, the concept of “systematic internalisers” was created. It was the legal form envisaged for banks that systematically execute their clients’ orders away from public markets. Other bilateral trading that was characterised as ad hoc, irregular, in large order sizes and between professional counterparties,[9] was considered to be an exception. As such, it was allowed to be traded under the label “over the counter” (OTC).

---

[6] EU Council [6].

[7] MIFID1 was the cornerstone of the Financial Services Action Plan (FSAP). The FSAP was an EU initiative running from 1999 for a period of 5 years. It aimed at improving the single market for financial services. See EU Commission [7].

[8] The purest form of the concentration rule pre-MiFID1 was found in France, Spain and Italy. Weaker forms—such as exchange precedence rule—were observed in Germany, whereas the UK had a de facto concentration rule in the form of obligation to report any off-exchange trade through the post-trade reporting facilities of the London Stock Exchange.

[9] MiFID1 (2004/39/EC), recital 53.

As a consequence of this liberalisation, a large number of trading venues emerged and challenged incumbent cash equity exchanges on their home turf. Within a few years, and after several rounds of industry consolidation, the equities market landscape in 2015 was formed by the home markets (major players include Bolsas y Mercados Españoles, Euronext, London Stock Exchange Group, Nasdaq OMX and Deutsche Börse) and by MTF platforms (major players are BATS Chi-X Europe[10] and LSEG-majority-owned Turquoise). As of today, the challengers capture around 20–40 % of market share in the respective home markets of the most relevant European indexes.

The flip side of competition is the fragmentation of liquidity. As liquidity is now dispersed across several trading venues, investors notice it is more difficult to find. This observation has two major adverse consequences, both which are related to the efficient reconsolidation of fragmented markets[11]:

1. Market participants need access to liquidity on dispersed trading venues. As a consequence, heavy investments in trading infrastructure were necessary.
2. The price formation process might suffer under certain conditions, especially when transparency is reduced.

With regard to the ***infrastructure investments***: Intermediaries such as banks and brokers need to have access to the dispersed liquidity pool(s) in order to enable best execution to their clients. On the one hand, they enjoy the reduced transaction costs that are the direct result of increased contestability between trading venues. On the other hand, brokers need to connect to various execution venues, resulting in heavy IT investments for themselves. It is still far from clear whether the savings in trading costs were able to overcompensate for the additional investments.

With regard to the ***price formation process***, the fathers of MiFID1 have considered that the adverse effects of competition could be healed by a large degree of transparency. If liquidity is dispersed, price formation can only work effectively if everyone can see the prices of orders that are offered on each market fragment. However, the requirement towards transparency usually clashes with the willingness of traders to provide information on their own orders. This is because when a large order gets exposed, prices tend to move unfavourably at the expense of the large order trader.[12] Most curiously, the very same traders that are unwilling to show their own orders are very keen to see the orders of everyone else. This has led to a situation where the public interest *to obtain transparency* clashes with the rational behaviour of individuals *not to provide transparency* themselves. This is where the regulator has to step in and mandate transparency. As MiFID1 introduced competition (i.e. the fragmentation of order flow), transparency was considered to be the tool to virtually reconsolidate (i.e. defragment) the market.

[10] Chi-X was acquired by BATS in 2011 and runs as "BATS Chi-X Europe" since. It has become a regulated investment exchange in May 2013.

[11] See Harris ([8], p. 533).

[12] For an overview about the reasons of the adverse price movements see, e.g., Harris ([8], p. 324–325).

Exceptions of the overall transparency requirement can be reasonable, for example to protect large orders from obtaining worse prices once their existence becomes known to the market ("adverse price movement"). Therefore, MiFID1 provided for some limited exceptions, which are known as "pre-trade transparency waivers" (see Sect. 25.3.2.1).

25.2.3 The Undesired Consequences of MiFID1

MiFID1 was written before the second wave of trading automation took off (see Sect. 25.2.1). As a result, the new technology developments were not appropriately considered in the legislation. These developments surfaced during the second half of the 2000s. For example, the phenomenon of algorithmic trading started to develop at the time, and the term high-frequency trading was coined around 2006.

MiFID1 did not offer the tools necessary to cope with these developments. As a result, ESMA issued Guidelines on Automated Trading[13] in 2012 introducing requirements for both trading venues and investment firms.

Another example includes brokers starting to automate the processes of their internal trading desks, resulting in the automation of the execution process. As a result, internal broker crossing networks (BCN) were developed, which allowed brokers to internalise their client orders more efficiently. BCNs also started to look more and more like exchanges—with the additional client benefit of not having to display their orders to the public. Orders executed in these BCNs were classified as "OTC" executions with the effect that the strict transparency obligations as imposed to trading venues did not apply. This was a straightforward case of regulatory arbitrage, as the principle "same business—same rules" did not apply to these exchange-like facilities.

A related phenomenon is the limited adoption of the "systematic internaliser" (SI) regime. Originally, it was intended to capture traditional OTC trading conducted between an investment firm and its client. However, the result was that only a handful of firms have applied for SI status[14]—a clear indication that the intended SI regime does not work as envisaged. In 2014, only 1 % of overall volumes in DAX equities were executed under the label "SI". This contrasts to as much as 43 % of volumes reported under the label of "OTC" (Fig. 25.2).

How can the large OTC portion be explained? One of the most prominent arguments used by the proponents of dark trading is that large orders need protection from the price moving unfavourably. However, due to the poor data quality (see Sect. 25.3.2.2) it is very difficult to verify this claim. In this context, one of the rare pieces of research[15] that have conducted a thorough analysis of the OTC market on a trade-by-trade basis has shown that an overwhelming part of orders executed in the OTC market is smaller than

---

[13] ESMA [9].

[14] A total of 12 firms were registered as SIs in the MiFID database as of January 2015: thereof 8 from the UK, 2 from Denmark and 1 each from France and Italy.

[15] Gomber, Lutat, Pierron and Weber [10].

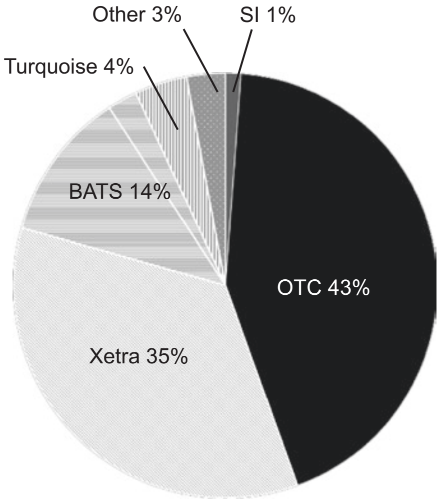


Fig. 25.2 Distribution of DAX trading between trading venues, SIs and OTC in 2014. Source: Fidessa, own calculations

the size available for the best price at the home market. This means that these orders could have been executed in the transparent market without any price impact. Thus the authors derive the finding that these orders neither need the "large order protection", nor do they fit the criteria for OTC trading mentioned in Sect. 25.2.2.

Similarly, the use of dark trading on trading venues was accelerated. By December 2014, around 4.4 % of total pan-European trading venue volume was conducted in regulated dark pools (i.e. a trading venue operating under a pre-trade transparency waiver).[16] The more attractive dark trading becomes, less activity takes place on transparent ("lit") venues, which may end up in the weakening of the price discovery process. Usually, the prices from the lit market are used as reference points in the dark market in a similar way to a beam from a lighthouse.[17] Therefore, all activities that weaken the price discovery process might have a detrimental effect on the collective wish for market transparency.

The MiFID revision provides a unique opportunity to address these issues. In our next section, we will take a closer look at these developments.

25.3 MiFID2: From Electronic Trading to Transparency and the Strengthening of the Financial System

MiFID1 contained a clause that mandated an assessment of various MiFID topics to take place by 2008. Due to the financial crisis, this undertaking was delayed until December 2010, when the public consultation for the review of MiFID1 was

[16] Source: Fidessa.

[17] Schwartz ([11], p. 340).

published.[18] In the following EU legislative procedure, the review consisted of a regulation and a directive and was developed in the period between October 2011[19] and June 2014.[20]

This section investigates the concrete changes brought by MiFID2. In the aftermath of the financial crisis, the aims have somewhat shifted from fostering efficiency and integrity towards increasing stability and strengthening the financial system.

25.3.1 Electronic Trading

One of the shortcomings of MIFID1 was that the technical developments that took place in the second part of the 2000s were not considered adequately. The MiFID revision provides this opportunity. As a consequence, MiFID2 provides significant provisions on algorithmic trading (AT) and HFT.

A detailed analysis of both AT and HFT is provided in Chap. 12 and in Chap. 13 of this volume, whereas the latter contribution also covers regulatory aspects.

As already mentioned in Sect. 25.2.3, one of the undesired consequences of MiFID with regard to electronic trading were BCNs. These systems have the capability to match client orders internally, without the need for the broker to pass them on to an external venue. In some cases, order flow from the proprietary trading desk was also mixed. These systems operate like quasi-exchanges, and in most cases without a market model nor trading surveillance. For these reasons, MiFID2 bans BCNs for equities. For other asset classes (i.e. bonds, structured finance products, emission allowances or derivatives), the concept of organised trading facilities (OTF) is introduced.[21]

25.3.2 Transparency

During the years of the MiFID review in the first half of the 2010s, transparency was one of the most controversially discussed topics in the industry. On the one hand there are the representatives of investors, who claimed that extensive order display would harm their execution costs; on the other hand the guardians of price discovery, claiming that almost every order must be displayed in order to contribute to price discovery in a fragmented environment.

MiFID1 has introduced provisions for equities transparency only. The aim of MiFID2 is twofold: First, it extends the transparency requirement for equities. Second, it introduces transparency requirements to other asset classes, which are not further considered within the scope of this chapter.

---

[18] EU Commission [12].

[19] EU Commission [13–15].

[20] EU Parliament and Council [4, 5].

[21] See MiFID2 (2014/65/EU) Article 4(1) (23) and Article 20.

With regard to equities trading, two types of transparency matter: transparency on orders (pre-trade transparency) and transparency on trades (post-trade transparency). In the following, the MiFID2 changes are presented.

25.3.2.1 Pre-trade Transparency

MiFID1 introduced a strict transparency regime for equities traded on trading venues. Exceptions were only allowed in four cases (so-called waivers):

- Reference price waiver (RPW)
- Negotiated trade waiver (NTW)
- Large-in-scale waiver
- Order management facility waiver

Throughout the legislative procedure, some of these waivers were intensely discussed, and there were proposals to drop some of them.[22] However, all of those waivers survived the MiFID review, but some with significant changes: The RPW and the NTW were restricted in their usage. Figure 25.3 shows the tremendous growth in the share of the RPW market, supporting the assumption that not tightening the requirements could lead to further substantial increases in dark trading.

The new rules for the RPW and the NTW state that trading under these can only be conducted as long as certain volume thresholds are not reached. These, so-called volume caps[23] amount to

- Four percent of volume traded on one dark venue compared to all trading venues in the EU or
- Eight percent of volume traded on all dark venues compared to all trading venues in the EU.

These figures are based on a rolling average over the last 12 months. Once the 4 % (8 %) cap is reached, the competent authority (all competent authorities) that authorised the use of the waivers by that venue has to suspend the use of that waiver on that venue (all venues) in that financial instrument for a period of 6 months.

Pre-trade transparency in equities and equity-like instruments by systematic internalisers is provided by their quotes. Title III of MiFIR specifies these requirements and mandates that SIs must display their quotes in liquid equities, which shall amount to at least 10 % of the standard market size (SMS). The SMS in a typical liquid German equity today amounts to 7.500 Euro.[24] As a result, SIs will be obliged

[22] EU Commission ([13], p. 8, and Article 4).

[23] See MiFIR, Article 5.

[24] According to MiFID Database (inquiry on 5 January 2015). Out of 111 shares that qualify the criteria of liquid German shares, the only exemptions are Allianz, Münchener Rück and Volkswagen, each with an SMS of 15,000 Euro.

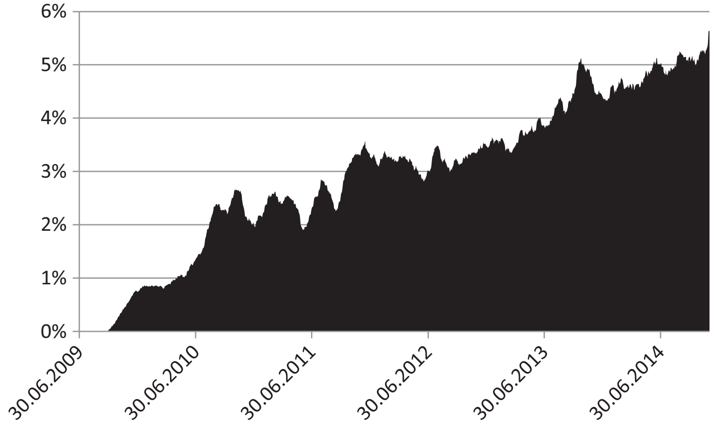


Fig. 25.3 DAX market share development of markets implementing the reference price waiver. Own analysis

to typically quote sizes as small as 750 Euro on each market side. In addition, SIs will be able to provide price improvement, provided that the prices fall within a public range close to market conditions.

As outlined in Sect. 25.2.3, only a handful of firms opted in to being an SI. This optionality was obviously seen as a flaw in MiFID1; therefore MIFIR introduced strict quantitative criteria, based on whether the firm provides "frequent and systematic" and "substantial" trading. Once the thresholds are crossed, the firm needs to register as SI in the respective financial instrument.

25.3.2.2 Post-trade Transparency

In Sect. 25.2.3 we have provided one possible explanation regarding the size of OTC trading: A vast number of small orders executed "OTC", although they could have been executed without price impact on the transparent market. Another explanation focuses on the quality of data reported under the label "OTC". Accordingly, there are no strict requirements towards how and when to report an OTC trade to the reporting venue. As a consequence, the same trade may (or may not) be reported several times (by both the buying broker and the selling broker; and then again by the two brokers trading with their respective end investors, resulting in up to four trade reports for one transaction). In addition, reports are often done with significant time delay. For example, MiFID1 allows the delay for up to 3 trading days, depending on the size of the order. As a result, OTC trade data is considered to be inflated, and inaccurate regarding the precise time of execution.

As a solution, MIFIR mandates investment firms to publish price, volume and time of their transactions through an approved publication arrangement (APA).[25] APAs are

[25] MiFIR Article 20.

facilities that filter post-trade data received from investment firms before publishing it. In order to further improve data quality, ESMA was mandated to produce technical standards specifying the maximum time limits allowed for publication, as well as the content in the form of specifying trade identifiers. These measures, combined with the publication of data through APAs, should ensure that data provided by investment firms is quality-checked, thereby ensuring a minimum informational content of bilateral trades.

25.3.3 Other Topics Relevant for Cash Equity Trading

In the following, we shed some light on additional topics that are noteworthy from a cash equities point of view as some significant changes are imminent.

25.3.3.1 Trading Obligation for Shares

As outlined in Sect. 25.2.3, the SI regime did not function effectively. MiFIR tackles the issue by demanding that more trading takes place on trading venues and SIs. As a result, MiFIR Article 23 introduces a trading obligation for investment firms, mandating that all trades in ***shares*** must be facilitated on a trading venue or by the investment firm in its capacity as SI. Exceptions apply for transactions that are non-systematic, ad hoc, irregular and infrequent or are carried out between eligible and professional counterparties and do not contribute to the price discovery process.

25.3.3.2 Transaction Reporting

Transaction reporting is the reporting of all trades to the competent authorities with the purpose to support the facilitation of their supervisory activities. This provision, introduced with MiFID1, related not only to equities, but also any financial instrument that is traded on an exchange. According to the provision, investment firms must report their trades to their competent authority at the next trading day at the latest.

Title IV MiFIR broadens the reporting requirements significantly. First, the scope of financial instruments that require reporting is increased. Second, the identity of the client on whose behalf the investment firms executed the transaction must be disclosed—with MiFID1, this provision was left to the discretion of each member state. Third, there is a new requirement to report whether the transaction resulted from a short sale. And finally, transaction reports will have to identify the persons (trader ID) and any computer algorithm (algo ID) that is responsible for the investment decision.

25.4 From Good Intentions to Possible Overregulation

In this section, we investigate the consequences of the new rules for implementation. We focus on the good intentions, and also highlight the dangers of overregulation from past regulatory initiatives involving cash equity. Then, we will focus on items from MiFID2 that seem not to be proportionate in terms of costs and benefits.

25.4.1 The Good Intentions of MiFID2

There are several things that the MiFID review has improved. In addition to safeguarding investor protection, achieving an efficient and integrated financial market, it was also important for regulators to position MiFID2 as "an essential part of the structural reforms aimed at establishing a safer, sounder, more transparent and more responsible financial system working for the economy and society as a whole in the aftermath of the financial crisis".[26] Another goal for MiFID2 was to address intransparent and unregulated parts of the financial system. This approach will definitely benefit the European market structure.

Some examples of how MiFID2 tackles these challenges are the following:

- Extension of transparency requirements for equities (see Sect. 25.3.2): Helps to preserve the price formation process on fragmented market.
- Introduction of transparency requirements for other asset classes, i.e. other asset classes, i.e. equity-likes, bonds, structured finance products, emission allowances or derivatives: Helps to establish the wider goal to trade more financial instruments in a transparent environment.
- Introduction of trading obligation for shares (see Sect. 25.3.3.1): Helps to ensure more trading on regulated execution venues.
- No OTF for equities (see Sect. 25.3.1): Prevents to water down the high market efficiency of the equity market, which would occur due to the ability for discretionary execution on this trading venue. In contrast, non-equities markets are not as developed, so that any kind of organised trading will help to increase the degree of efficiency on these market segments.

Despite these valid regulatory actions, there is always the latent danger of overregulation. This holds especially true in a situation where one of the newly issued goals of the lawmaker explicitly considers the financial crisis.

[26] EU Commission ([12], p. 6).

25.4.2 The Pitfalls of Overregulation

The constant challenge for lawmakers and regulators is to find the right balance between prudent regulation and overregulation. Quite a few regulatory initiatives have been issued in the aftermath of the financial crisis of 2008, MiFID2 being one of them. There are some examples from past experience to consider when drafting new regulations, as there have been pitfalls which were in the most cases successfully avoided:

- Short-selling rules: During the crisis and in its aftermath, a series of European short-selling rules have been crafted, culminating in the EU Short-Selling Regulation (SSR).[27] All these rules had in common the restriction of short-selling activities, especially the ban of uncovered short selling. But short selling is essential for liquidity provisioning strategies. Therefore, the good intention of regulators to address market stability is in reality fully reversed and the eventual outcome is lower liquidity and higher trading costs for investors. Initial proposals included the ban of uncovered short positions (including derivatives). This would have restricted the ability of market makers to perform their duties. Fortunately, those plans were shelved and the final version of the Regulation included exemptions for intraday uncovered short positions for liquid equities, fostering market efficiency.
- Rules regarding algorithmic trading and HFT: In the aftermath of the US Flash Crash of 2010, automated trading got under the scrutiny of regulators. The key question is whether entrusting trading decisions to machines would endanger market stability and execution process. As a result, ESMA published its Guidelines on Automated Trading,[28] and Germany passed a law that governs automated trading activity.[29] By and large, the provisions of the German HFT Act carefully balance the risks and benefits of automated trading and re-considered to be proportionate and reasonable. Most of these provisions can in a similar fashion also be found in MIFID2. However, during the lawmaking process, some less helpful proposals affecting the market structure have been made, for instance the idea to introduce a mandatory minimum order resting time of half a second for all orders. Such an idea would have had a disruptive effect on the market design, as the currently proved and functioning market models of trading venues would have to be changed, whereas these have neither been thoroughly thought through nor properly tested in practice before.
- The planned introduction of a financial transaction tax (FTT): The idea of the FTT is often regarded as panacea to "making the financial sector pay its fair share"[30] for the role it played in the origins of the economic crisis, with the losses

[27] EU Parliament and Council [16].

[28] ESMA [9].

[29] Roth and Budimir [17].

[30] EU Commission [14].

eventually imposed to taxpayers. But in the end, an FTT in equities will simply shift trading volumes to the dark and unregulated execution space, and will most certainly increase the cost of investing to ordinary investors.

In the next section, we discuss what undesired consequences might result from MiFID2.

25.4.3 Possible Overregulation with MiFID2

In the following, we will highlight some points where the efforts might exceed the prospected benefits of the regulation.

25.4.3.1 Electronic Trading

MIFID1 implemented several exemptions for various market participants not falling into its scope. One such exemption existed for pure proprietary traders, with the result that HFT were not in the scope of MiFID. This changes with MiFID2, which envisages that all proprietary trading firms engaging in direct electronic access or entities applying a high-frequency algorithmic technique fall under its scope. This effectively introduces licensing requirements for these participants and thereby increases the ability for regulators to perform their supervisory activities.

On the flip side, for the relevant firms this also increases their regulatory burden. Licensing the status might lead to a situation where proprietary trading firms need to satisfy numerous additional requirements imposed by their local regulations. For example, these requirements might envisage that becoming a regulated investment firm goes along with mandatory contributions towards deposit insurance funds—although HFT firms per definition do not have customers with funds that need to be safeguarded. Eventually, the strengthened requirements might lead to a situation where these firms adapt their business models and cease trading with the overall result of ceasing to provide liquidity to the market, with the result of increased spreads.

MiFID2 introduces the requirement to flag trading algorithms. The purpose of this provision is to equip the supervisor with the ability to distinguish individual strategies that reside behind a trading participant. The algo flagging requirement is stated in the Directive (2014/65/EU, Article 48(10)), with the consequence that each member state might impose its own technical details on the flagging requirements. As the institutional trading landscape in Europe is highly integrated, and most of the participants trade on all European venues, a heterogeneous flagging requirement across member states would highly increase the complexity of flagging. In such a scenario, any participant would have to apply a different flagging logic, depending on the country of origin of the respective trading venue that the participant acts on. A far better solution is provided by a harmonisation of flagging requirements throughout Europe. An opportunity to achieve this is provided by the transaction reporting regime as outlined by MiFIR (see Sect. 25.4.3.3).

25.4.3.2 Transparency

MiFID2 has introduced volume caps for waivers, as outlined in Sect. 25.3.2.1. These caps will require data processing on a daily basis for every venue that is engaged in dark trading (RPW and NTW). For each financial instrument, the data has to be obtained, calculated, cross-checked and then evaluated whether the thresholds are breached. In the case of the 4 or 8 % volume cap breach, pan-European coordination is necessary to shut the respective markets implementing this waiver. This will require a high coordination effort, involving many parties. In addition to the increased alignment efforts, there is also the question of liability when things do not develop as planned.

The desired result—the protection of market transparency—could have been achieved way more elegantly by an alternate approach: Instead of capping the total amount of trading (often achieved by trading very small undisplayed orders), a simple order size threshold would have achieved the same goal: For instance, an order volume threshold set at a percentage (e.g. in the area of 50–80 %) of the current large-in-scale threshold would have enabled only larger orders to be traded by this waiver. Smaller orders—for which there is no need for protection from display as their price impact is small—would have continued to be executed in the lit order book environment, thereby positively contributing to price discovery.

25.4.3.3 Other Aspects

The trading obligation makes sure that shares are by and large executed on a trading venue or via an SI. This would mean that it would become impossible to continue OTC share trading on a systematic and frequent basis. However, one market structure parameter was overlooked that enables for competition: tick size. As tick sizes are regulated in MiFID2 across trading venues,[31] there is no possibility for exchanges and MTFs to compete against each other on this parameter alone. But tick sizes are not regulated for systematic internalisers. Therefore it is expected that SIs will compete with trading venues on tick sizes, offering slightly better prices to their clients than that are possible on trading venues.

With regard to transaction reporting, MiFIR requires to report whether the transaction resulted from a short sale. The definition of short selling is taken from the SSR, which also requires the reporting of short sales to the regulator and the public. However, these reports require only net positions to be reported. In contrast, MiFIR requires to designate every trade whether it resulted from a short sale or not. This

[31] MiFID2 (2014/65/EU) Article 49.

will increase the burden on the back offices of trading firms, as they will have to manage their inventory very thoroughly in order to be compliant to this provision. This will introduce a further complexity, which will most likely increase infrastructure costs. Behind the background of the existing short-selling transparency regime, additional insights for regulators seem to be limited whereas the burden of the implementation might be considerable.

MiFID2 has implemented provisions for flagging of trading algorithms. However, the legislative text did not provide for a link between the regulation and the directive. As outlined in Sect. 25.4.3.1, MiFID2 (2014/65/EU) allows member states to require different methods of algo flagging, whereas MiFIR (Article 26(3)) demands a unique, pan-European solution. The obvious pitfall in this setting is that due to lack of coordination between the EU and the member states the implementation might not be standardised, resulting in increased efforts for participants, trading venues and competent authorities.

References

1. EU Parliament and Council (2004): Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004 on markets in financial instruments amending Council Directives 85/611/EEC and 93/6/EEC and Directive 2000/12/EC of the European Parliament and the Council and Repealing Council Directive 93/22/EEC. Official Journal of the European Union. L 145/1–61, 30. April 2004.
2. EU Commission (2006): Commission Regulation (EC) No. 1287/2006 of 10 August 2006 implementing Directive 2004/39/EC of the European Parliament and of the Council as regards record-keeping obligations for investment firms, transaction reporting, market transparency, admission of financial instruments to trading, and defined terms for the purposes of that Directive. Official Journal of the European Union. L 241/1–25.
3. EU Commission (2006): Commission Directive 2006/73/EC of 10 August 2006 implementing Directive 2004/39/EC of the European Parliament and of the Council as Regards Organisational Requirements and Operating Conditions for Investment Firms and Defined Terms for the Purposes of that Directive. L 241/26–58.
4. EU Parliament and Council (2014): "Regulation (EU) No 600/2014 of the European Parliament and of the Council of 15 May 2014 on Markets in Financial Instruments and Amending Regulation EU No 648/2012." Official Journal of the European Union. Volume 57, 12 June 2014, L173/84–148.
5. EU Parliament and Council (2014): "Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on Markets in Financial Instruments and Amending Directive 2002/92/EC and Directive 2011/61/EU (recast)." Official Journal of the European Union. Volume 57, 12 June 2014, L173/349-469.
6. EU Council (1993): "Council Directive 93/22/EEC of 10 May 1993 on investment services in the securities field." Official Journal of the European Communities. L141, Volume 36, 11 June 1993.
7. EU Commission (2009): Financial Services Action Plan. Available at http://ec.europa.eu/finance/general-policy/actionplan/index_en.htm
8. Harris, Larry (2003): Trading and Exchanges: Market Microstructure for Practitioners. Oxford University Press. Oxford, New York.

9. ESMA (2012): Guidelines: Systems and controls in an automated trading environment for trading platforms, investment firms and competent authorities. European Securities and Markets Authority. 24 February 2012. ESMA/2012/122 (EN).
10. P. Gomber, M. Lutat, A. Pierron and M. Christian Weber (2011): Shedding Light on the Dark—OTC Equities Trading in Europe. Journal of Trading, 6(1): 74–86.
11. Schwartz, Robert A. (2010): Micro Markets: A Market Structure Approach to Microeconomic Analysis. Hoboken, NJ, John Wiley & Sons.
12. EU Commission (2010): Public Consultation—Review of the Markets in Financial Instruments Directive (MiFID). 8 December 2010.
13. EU Commission (2011): "New rules for more efficient, resilient and transparent financial markets in Europe". Press Release, 20. October 2011.
14. EU Commission (2011): Proposal for a Regulation of the European Parliament and of the Concil on Markets in Financial Instruments and Amending Regulation [EMIR] on OTC Derivatives, Central Counterparties and Trade Repositories. COM(2011), 652 final, 2011/0296 (COD), 20. October 2011.
15. EU Commission (2011): Financial Transaction Tax: Making the Financial Sector Pay its fair Share. Press Release. 28 September 2011.
16. EU Parliament and Council (2012): Regulation (EU) No. 236/2012 of the European Parliament and of the Council of 14 March 2012 on Short Selling and Certain aspects of Credit Default Swaps. Official Journal of the European Union. L86/1-24.
17. Roth, Randolf and Miroslav Budimir (2014): Regulating Automated Trading: Eurex's Response to the German HFT Act. Futures Industry. 24(1):38–40.

Further Reading

EU Commission website on MiFID 1 and MiFID 2: http://ec.europa.eu/finance/securities/isd/index_en.htm

Foresight (2012): Economic impact assessments on MiFID II policy measures related to computer trading in financial markets. Foresight, UK Government Office for Science Working Paper. Ref: 12/1088

Gomber, Peter and Beneidkt Jaeger (2014): MiFID: Eine systematische Analyse der Zielerreichung. In: Zeitschrift fuer Bankrecht und Bankwirtschaft, 2014 No. 1, pp. 40–53; 2014

Chapter 26
European Financial Integration: Monetary Union, Banking Union, Capital Markets Union

Andreas Dombret

26.1 Introduction

Since the Treaties of Rome were concluded in 1957, the history of Europe has been marked by an ever-deepening integration. To some degree this process has certainly been driven by the idea that in a globalised world only a united and strong Europe can succeed. In 1954, Jean Monnet, one of the founding fathers of the European Union, said: "Our countries have become too small for today's world, when compared to the potential of modern technical means and in relation to the dimension of America and Russia today, China and India tomorrow [1]". In that sense, European integration and also the introduction of the euro could be interpreted as a response to globalisation—as an attempt to create a strong regional pole in an increasingly multipolar world.

From an economic point of view, this process of integration culminated in the introduction of the euro in 1999. In that year, Europe took a historical step towards deeper financial integration. However, it was a controversial step. Many critics argued that Europe was not yet an optimal currency area and thus not ready for a single currency. In their "coronation theory", the creation of a single currency should be the coronation of a long-term process of political and economic convergence, not the precondition for it. In contrast to that view, proponents of the euro propagated the "locomotive theory" whereby a single currency would be introduced early on and would act as a locomotive for further economic and political integration.[1] Eventually, the proponents of the "locomotive theory" prevailed, and the euro was introduced in 1999 and, for the first 10 years of its existence, proved to be a stable currency.

[1] For a more extensive discussion on these two theories see Kruse, D. C. [2], Monetary Integration in Western Europe: EMU, EMS and Beyond. London, and Wolf, D. [3], Neofuctionalism and Intergovernmentalism Amalgamated: The Case of EMU. In: Verdun, A., The Euro: European Integration Theory and Economic and Monetary Union. New York.

A. Dombret (✉)
Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany
e-mail: andreas.dombret@bundesbank.de

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_26

Then, in the wake of the global financial crisis of 2008, the euro area slid into a crisis of its own. In 2010, Greece stumbled into a sovereign debt crisis which quickly led to a loss of confidence in other countries at the periphery of the euro area and eventually brought the euro area to the brink of collapse. Extensive rescue packages provided by the member states of the euro area as well as unconventional measures taken by the ECB helped to calm the markets and prevented the crisis from escalating. At the time of writing, many euro-area countries are still struggling with the fallout from the crisis such as high unemployment, sluggish growth and high public debt. Some of the necessary reforms at the national level have already been undertaken and are beginning to bear fruit, but a long way still lies ahead.

At some point along that way, the focus should shift from managing the current crisis to preventing future crises. And here it is relevant to look at the framework of monetary union. The crisis has exposed weaknesses in this framework, and looking to the future, it is these weaknesses we have to address. In the attempt to reach this objective, integration could be the way ahead. But before discussing the way that lies before us, it is important to look back in order to understand how it all happened.

26.2 The European Monetary Union in Crisis

The crisis in the euro area is a complex phenomenon that can be viewed from a great many different angles. With regard to the institutional framework of monetary union, there are at least two aspects that deserve to be highlighted: first, a fateful connection between banks and public finances; second, a somewhat lopsided approach towards European integration that resulted in an imbalance between liability and control.

26.2.1 Banks, Public Finances and the "Doom Loop"

A central aspect of the crisis in the euro area was the connection between banks and public finances. This connection can take the form of a vicious circle—or doom loop—and was a major driver of the crisis. This doom loop is to a large extent driven by a concept that is sometimes labelled "systemic relevance" and sometimes "too big to fail". What does it mean when a bank is systemically relevant or "too big to fail"? In essence, it means that the bank in question is very big or very interconnected. It is woven so deeply into the fabric of the financial system that its failure might cause a rupture of the entire system. The most prominent example is the fall of Lehman Brothers in September 2008, which marked the beginning of the global financial crisis.

The consequence of "too big to fail" or systemic relevance is that whenever a big or interconnected bank gets into trouble, the government might be compelled to bail it out in order to avoid a full-blown financial crisis. This means that banks that are "too big to fail" have an implicit and cost-free insurance: whatever happens, the government is likely to come to their aid. And that is what happened when the financial crisis unfolded in 2008 and banks everywhere ran into difficulties. Governments had to step in to avoid a breakdown of the entire financial system. However, saving banks is a costly endeavour for governments. Ireland is a case in point. In order to save its banking sector, Ireland ran up a budget deficit amounting to 30 % of its economic output in 2010. In the same year, Germany added almost €33 billion to its public deficit as a consequence of supporting the banking system.

Obviously, saving banks puts a lot of strain on public finances, which directly relates to the other side of the doom loop. When public finances run into difficulties, the banks are put under strain. Fiscal distress is particularly problematic for those banks that hold government bonds. Thus, a simultaneous crisis in the banking system and in public finances can emerge that is driven by a self-propagating mechanism. This is what happened in the euro area.

26.2.2 Liability, Control and Lopsided Integration

In addition to the doom loop, the crisis in the euro area also exposed flaws in the institutional framework of monetary union. To understand these flaws, it is important to be familiar with the particular features of the European monetary union. The European monetary union is special in that it combines a single monetary policy with national fiscal policies. The monetary policy for the 18 countries of the euro area is decided by the Governing Council of the ECB in Frankfurt. However, the fiscal policies of the 18 euro-area member states are a matter for the national policy makers—each country decides on its own government revenues and expenditures.

This imbalance of liability and control gives individual countries an incentive to borrow—a "deficit bias" is built into the system. The costs of borrowing are spread across all the member states of the monetary union—for example, by means of a higher interest rate level for all of them. And in the extreme case of a sovereign debt crisis, the costs for the other members of monetary union would be very high, as they would be liable for the decisions of one individual country. For this reason, this imbalance of liability and control encourages the member states to borrow more than they would if they had their own currency. This "deficit bias" was also recognised by the founders of the monetary union. In order to reduce it, they did two things. First, they created explicit rules on borrowing in the form of the Stability and Growth Pact. This pact was intended to keep a tight check on national fiscal policies. Second, the founders of the monetary union incorporated the "no bail-out" principle into the Maastricht Treaty: no euro-area country was to be liable for the debts of another member state. Thus, individual responsibility was to be the guiding principle for fiscal policy in the monetary union in that each country was itself to bear the consequences of its own fiscal policy.

The rules of the Stability and Growth Pact were intended to keep borrowing by the euro-area countries within reasonable limits. But rules were not the only means of achieving this. The financial market actors, too, had to ensure that the euro-area countries did not incur excessive debt. The idea behind this is simple: if a country were to become excessively indebted, markets would demand higher risk premia when lending money to this particular country. For the country in question, then, borrowing would become more expensive, and it would have to reduce its debt to a sustainable level. However, neither of the two safeguards worked. Neither the discipline of the financial markets nor the rules were able to prevent individual countries from running up excessive debt. Investors on the financial markets did not differentiate between individual countries after the euro was introduced. Risk premia across the euro area converged at a very low level. Consequently, the markets tolerated the problems of individual countries for far too long. At the same time, policy makers stretched and sometimes ignored the rules of the Stability and Growth Pact.

Then, in 2008, the financial crisis broke out. Subsequently, many countries had to rescue their banking systems and support economic activity. And that drove up their levels of sovereign debt dramatically. At the same time, investors on the financial markets suddenly seemed to become aware of the problems which some countries were experiencing. Now they saw the high level of sovereign debt, the lack of competitiveness and the risk of contagion effects between the individual countries as well as between banks and public finances. In short, they lost confidence in the crisis-hit countries. But this also meant that capital flows dried up—capital flows that had previously covered up all the problems.

26.2.3 Fiscal Union, Banking Union and Capital Markets Union?

This chapter has highlighted two structural weaknesses of the monetary union framework that contributed to the recent crisis. To be sure, the crisis is a complex phenomenon with an equally complex and broad range of sources. However, the two issues discussed here provide policy makers with a rather clear agenda on how to put monetary union on a more solid footing. The necessary steps all point in the direction of deeper integration. In the following chapters three areas are discussed in which deeper integration could contribute to make monetary union more stable: fiscal policy, banking supervision and capital markets.

26.3 Fiscal Union: Beyond the Horizon

In Sect. 26.2.2, lopsided integration and the resulting imbalance between liability and control were identified as a central weakness of the institutional framework of monetary union. This imbalance has been increased by those measures that were taken in order to contain the crisis. In response to the crisis, rescue mechanisms such as the European Financial Stability Facility (EFSF) and the European Stability Mechanism (ESM) were put in place, and they did manage to contain the fallout from the crisis somewhat. But at the same time, these rescue mechanisms have further weakened the principle of individual responsibility: fiscal responsibility has essentially remained national, while liabilities have been partially mutualised in a very direct fashion.

To put monetary union on a more solid footing, liability and control have to be rebalanced. One option to achieve this objective would be to deepen fiscal integration in the euro area. If we were to take this path, the European level would gain certain control rights over national budgets. This would amount to what is known as a fiscal union. However, such a step would depend on the countries of the euro area transferring national sovereignty to the European level—for example, by giving the European level the right to intervene in the event of unsound public finances. Giving up sovereignty in this way would be a radical change and require wide-ranging legislative changes nationally and at the European level. More than anything, such changes would need the support not only of policy makers but also of the general public. And on this point we need to be realistic. At present, there seems to be no willingness to do that—not in Germany or in any other country of the euro area.

This means that, for the foreseeable future, control of fiscal policy in Europe will remain at the national level. Thus, in this area, deeper integration still lies beyond the horizon. But if the current state of lopsided integration prevails, the national level must assume liability for its policies. That would mean strengthening the current set of rules on borrowing: the Stability and Growth Pact. These rules have since been tightened. However, past experience casts some doubt on the idea that stiffening the rules alone will suffice to enforce the principle of individual responsibility and rebalance liability and control. In essence, a balance between liability and control requires that sovereigns, banks and investors bear the consequences of their decisions. This implies that it is primarily up to the respective government and its citizens to come up with the funds required for repaying public debt. At the same time, it implies that the risk of non-repayment ultimately lies with the investors, since they are the ones who reap the return when things go well. And if the fiscal limit has been reached for real, public debt needs to be restructured without posing a systemic threat to financial stability.

The introduction of collective action clauses into sovereign bonds was a first step in that direction. But more steps are needed. The Bundesbank has put forward a proposal for sovereign bonds to include an automatic maturity extension of 3 years in the event that a sovereign accesses the European rescue mechanisms [4]. This automatic maturity extension would allow the sovereign in question to tackle its fiscal challenges while preventing investors from bolting. Liability and control would be rebalanced to some degree. The amount of official financial support would be reduced, and time would be bought to figure out if the problem is one of temporary illiquidity or outright insolvency. But ultimately, all these questions boil down to the quip of American economist Allan Meltzer: "Capitalism without failure is like religion without sin. It doesn't work". Thus, it has to be ensured that restructuring sovereign debt is possible without bringing down the financial system as a whole. The same goes for banks, and leads directly to the next area in which further integration could be the way forwards [5].

26.4 Banking Union: A New World

While fiscal union remains a rather unrealistic vision, another step towards integration has just been taken. On 4 November 2014, the ECB assumed direct supervision of the 120 largest banks in the euro area, thereby erecting the first pillar of a European banking union. The 120 banks which are now under the supervision of the ECB account for more than 80 % of the aggregated balance sheet for the euro-area banking sector. This has made the ECB one of the world's largest supervisors.

Implementing European banking supervision is certainly the biggest step towards financial integration in Europe since the launch of our common currency. And it is the most logical step to take. Single monetary policy requires integrated financial markets—which includes, without doubt, European-level banking supervision. European banking supervision will allow banks throughout the euro area to be supervised according to the same high standards. In addition, cross-border effects can be covered better through joint supervision than by national supervisors alone. Also, adding a European perspective to the national view will put more distance between the supervisory authority and the entities it supervises. This will minimise the danger of supervisors getting all too close to their banks and thus treating them with "kid gloves" out of national interest.

And, equally important, European banking supervision will contribute to cutting the doom loop by making the banking sector more stable and crises less likely. However, the problem of "too big to fail" remains. There are still banks that are so large that their failure could disrupt the whole system, and would therefore have to be saved by the government. Thus, European banking supervision has to be supplemented by a European mechanism that allows banks to fail without disrupting the system as a whole. From 2016 onwards, such a European resolution mechanism will be in place as the second pillar of European banking union. From then on, the rule will be that if a bank is no longer viable, shareholders and creditors will be first in line to bear that bank's losses, and taxpayers' money will only be the very last resort. Thus, the combination of European banking supervision and a European resolution mechanism will bring us much closer towards the goal of disentangling the close connection between banks and public finances.

26.5 Capital Markets Union: Tomorrow's Reality?

The banking union is definitely a major step forwards in designing a better framework for the European monetary union. However, we should broaden our view beyond the banking sector. A deeper integration of capital markets would also contribute to sharing opportunities and risks. To be sure, we have come a long way in integrating capital markets in Europe. According to statistics provided by the Bank for International Settlements, European banks' claims within Europe stood at 36 % of GDP at the end of the 1990s, growing to 77 % by 2008. This share has fallen during the crisis, but still stands at around 48 % of GDP.

However, there are two caveats regarding this trend of capital market integration in Europe. First, the financing structure of European companies is still predominantly bank based. A look at the balance sheets of German companies, for instance, shows that bank credit still accounts for about 15 % of the liability side. This is clearly lower than the 22 % observed at the end of the 1990s, but compared to the USA or the UK there is still room to increase the share of capital market financing. Second, although banks' cross-border exposures have risen, capital market integration remains incomplete. In the banking sector, for instance, integration has concentrated on the interbank market while credit markets for companies remain predominantly national. The integration of the capital markets may have increased in Europe, but the ownership structures of many companies have not. They are still strongly national.

The relatively low level of integration in European capital markets represents a barrier for risk sharing. Equity holdings in the USA, for instance, are much more widely dispersed throughout the entire country. Thus, when a negative shock hits an industry or a specific region, the resulting losses are spread widely beyond that region. In Europe they are not, because equity holdings are much more concentrated at the national level. Empirical studies for the USA show that integrated markets for capital cushion around 40 % of the cyclical fluctuations between the US federal states. A share of around 25 % is smoothed via the credit markets, while fiscal mechanisms cushion just 20 % of shocks [6, 7]. Studies for Germany come to similar conclusions [8].

Against this backdrop, two general lines of action could be followed in Europe. First, it might be beneficial to increase the share of capital markets in companies' financing structure. This would, of course, require a shift away from the traditional bank-based system to a certain degree. In this context, it might be worth taking a closer look at tax regimes, among other things. Currently, tax treatment still favours debt financing over equity financing. Removing this bias in taxation would encourage companies to strengthen their equity base and thus turn more towards capital markets in their search for sources of funds. The second line of action would be to deepen the integration of capital markets, which might eventually result in the formation of a capital markets union. To be sure, the concept of a capital markets union is not as clear-cut as that of a banking union. Capital markets are complex, and non-bank finance takes many forms: corporate bonds, private equity, public equity, venture capital or peer-to-peer lending, to name just a few. And integration relates not just to financial products but also to integral elements of the respective markets, such as stock exchanges and central counterparties. Thus, any attempt to form a capital markets union would require many different measures in many different areas [9].

Nevertheless, at the time of writing, the idea of a capital markets union is gaining some traction at the political level.[2] Indeed, it would be a logical step to supplement the banking union with a capital markets union. It would reduce fragmentation in European financial markets and, at the same time, enhance their efficiency and stability. Thus, it is certainly a goal worth pursuing.

26.6 Conclusion

George Washington is credited with having prophesised more than two centuries ago that a United States of Europe would come into being [10]. This is a bold vision which has been voiced repeatedly since Washington's days. However, the vision of a United States of Europe is a very broad approach aiming at an encompassing political integration.

This contribution has taken a more modest approach. Arguing from an economic standpoint, it has highlighted three areas in which deeper integration might help to enhance the stability of monetary union. The first area is public finances, although a fiscal union is currently a rather unrealistic vision. The second is the banking system, and here a major step towards integration has been taken with the European banking union. And the third is capital markets. Looking to the future, a capital markets union is another project that would contribute to enhancing the stability of monetary union.

To be sure, these are all big steps, but they are worth taking. A stable monetary union will eventually benefit all member states, including Germany.

References

1. Monnet, Jean (1978): Memoirs. London: Collins.
2. Kruse, D. C. (1980). Monetary Integration in Western Europe: EMU, EMS and Beyond. London.
3. Wolf, D. (2002). Neofuctionalism and Intergovernmentalism Amalgamated: The Case of EMU. In: Verdun, A. (Ed.), *The Euro: European Integration Theory and Economic and Monetary Union*. New York.
4. Weber, A., Ulbrich, J. and Wendorff, K. (2011). Krisenhilfe ohne Gemeinschaftshaftung. Frankfurter Allgemeine Zeitung, 3 March 2011.
5. Weidman, J. (2014). Reforms for Recovery and Resilience. Speech at Bank of Latvia Conference 2014, 17 October 2014.
6. Asdrubali P., Sørensen, B.E. and Yosha, O. (1996). Channels of Interstate Risk Sharing: US 1963–1990. *Quarterly Journal of Economics, vol. 111*(4), pp. 1081–1110.
7. Sørensen, B.E. and Yosha, O. (1998). International risk sharing and European monetary unification. *Journal of International Economics, vol. 45*(2), pp. 211–238.

[2] Among others, the president of the European Commission, Jean-Claude Juncker, has proposed a capital markets union at various occasions in 2014.

8. Hepp, R. and von Hagen, J. (2013). Interstate risk sharing in Germany: 1970–2006. *Oxford Economic Papers, 65*(1), 1–24.
9. Dombret, A. (2014). Designing a stable monetary union—Progress and open issues. Speech at The Brookings Institution, 8 October 2014.
10. Fulbright, J.W. (1948). A United States of Europe?. *Annals of the American Academy of Political and Social Science, vol. 257*, pp. 151–156.

Glossary

Stephanie Welkoborsky

Note: The Glossary pools four different sources for definitions in order to guarantee the highest possible level of neutrality. Barron's Dictionary of Finance and Investment Terms provides general market definitions, while the exchange's view is represented by official Deutsche Börse Group definitions and the industry perspective by excerpts of the UBS Banking Glossary and other leading sources. Where applicable, official documents have been cited to include the regulatory view.

S. Welkoborsky
Genossenschaftsverband e.V., Frankfurt, Germany

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9

Table A.1 Glossary

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Algorithmic Trading		Also: algo trading. Trading technique that involves the automatic output of buy and sell orders to the market by the participants' systems when predefined parameters are matched		
Alternative trading systems (ATS)	Electronic communication network: any one of a number of electronic systems that displays and matches orders placed on exchanges and over the counter by market makers and traders			1. Alternative trading system means any organization, association, person, group of persons, or system: (a) That constitutes, maintains, or provides a market place or facilities for bringing together purchasers and sellers of securities or for otherwise performing with respect to securities the functions commonly performed by a stock exchange within the meaning of § 240.3b-16 of this chapter; and (b) That does not: Set rules governing the conduct of subscribers other than the conduct of such subscribers' trading on such organization, association, person, group of persons, or system; or Discipline subscribers other than by exclusion from trading (17 CFR 242.300).

Asset servicing			Administration services provided by a central securities depository (CSD) or custodian in connection with the custody and/or safekeeping of financial instruments (e.g., the processing of corporate events or the handling of taxes)[f]	
Asset-backed security	Bonds or notes backed by loan paper or accounts receivable originated by banks, credit card companies, or other providers of credit and often "enhanced" by a bank letter of credit or by insurance coverage provided by an institution other than the issuer. Typically, the originator of the loan or accounts receivable paper sells it to a specially created trust, which repackages it as securities with a minimum denomination of $ 1000 a term of 5 years or less. The securities are then underwritten by brokerage firms who reoffer them to the public […]			The term "asset-backed security" means a fixed-income or other security collateralized by any type of self-liquidating financial asset (including a loan, a lease, a mortgage, or a secured or unsecured receivable) that allows the holder of the security to receive payments that depend primarily on cash flow from the asset (H. R. 4173—515).
Auction (market)	System by which securities are bought and sold through brokers on the securities exchanges, as distinguished from the over-the-counter market, where trades are negotiated […]. As in any auction, a price is established by competitive bidding between broker acting as agents for buyers and sellers		Public sale in which goods or services are sold or bought in a bidding process[e]	
Availability		In computer systems and networking, availability is the probability that a system will work as required when required over a specified period of time		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Back end		The central system components owned by the exchange which hosts the core functionality of the exchange trading system. The order books and the matching engine, for example, are essential back end components		
Back office	Bank or brokerage house departments not directly involved in selling or trading. The bank office sees to accounting records, compliance with government regulations, and communication between branches	In a financial services company, the back office in a trading environment carries out administration and support activities, such as record keeping, accounting, settlement, and regulatory compliance		
Backbone connections		High availability and high-bandwidth data network connections linking back-end components of the exchange trading system to the access points		
Bandwidth		The throughput capacity of a data transmission network component. Bandwidth is typically measured as data volume per second, e.g., in Megabit per second		
Basel III		Recommendations by the Basel Committee on Banking Supervision at the Bank for International Settlements in Basel, Switzerland. The aim of the recommendations is to ensure the stability of the financial system. They update and supplement the regulatory framework for banks (Basel II recommendations) that were resolved in 2004, especially to eliminate weaknesses of the framework which got visible during the global financial and economic crisis[c]		
Bearer security		Bearer shares differ from registered shares in that the holder of a bearer share is not designated by name on the share certificate, and is usually not required to furnish proof of rightful ownership. Bearer shares are transferred informally and by delivery, without any changes having to be made to the certificate. As a consequence, they are highly fungible (interchangeable) and can be traded easily[a]	A security having no facility for the issuer to record the owner of the security, and where proof of ownership is physical possession of the security certificate; historically, mainly bonds rather than equities were issued in bearer form. See registered security[h]	

Big data		A trend in information technology and financial services to retain and analyze massive amounts of client, market, and transaction data in large electronic databases, also called data warehouses		
Blue chip	Common stock of a nationally known company that has a long record of profit growth and/or dividend payment and a reputation for quality management, products and services [...]. Blue chip stocks typically are relatively high priced and have moderate dividend yields	Shares with large market capitalization, typically included in prominent indices, such as DAX[a]	Share of a company of prime borrower standing. Original meaning of word: fragment of a diamond[e]	
Book building		The process by which an underwriter attempts to determine at what price to offer an initial public offering of a company based on demand from investors	Also: book building procedure. Procedure in an issue price is set. In contrast to a fixed-price procedure, the issuer carries the placement risk for book building[e]	
Book entry		A method whereby transfer of ownership of securities is effected by debits and credits to accounts without the need for the movement of physical certificates or documents[d]		
Broker	Person who acts as an intermediary between a buyer and seller, usually charging a commission. A broker who is specialized in stocks, bonds, commodities, or options acts as agent and must be registered with the exchange where the securities are traded		Stockbroker: also: stock exchange dealer, broker. An individual or firm which trades professionally on the securities trader[e]	
Central bank liquidity		The process by which an underwriter attempts to determine at what price to offer an initial public offering of a company based on demand from investors[c]		
Central bank money			Liabilities of a central bank, in the form of either banknotes or bank deposits held at a central bank, which can be used for settlement purposes[f]	

(continued)

Term	Barron’s	Deutsche Börse group	Third party	Regulatory documents
Central counterparty (CCP)		Legal entity that acts as an intermediary between the parties to a securities or derivatives trade and is the seller to every buyer and the buyer to every seller, replacing the default risk of the original counterparty with its own and facilitating netting. Many CCPs also provide various other benefits, including post-trade anonymity, reporting, and risk management tools to their members[b]	Also: clearing organization. Institution which settles the payment of financial transactions between members. Used to rationalize securities trading and bank payments. Local clearing houses exist in all major banking centres, often alongside other organizations working on a national or international basis[e]	“CCP” means a legal person that interposes itself between the counterparties to the contracts traded on one or more financial markets, becoming the buyer to every seller and the seller to every buyer. Regulation (EU) No 648/2012 Art. 2 (1)
Central Securities Depositories Regulation		Designed to harmonize the European CSD legal framework, structure and services. It aims to introduce overall post-trade market discipline. Its impact will be widespread, as it will complement MIFID and EMIR[c]		
Central securities depository (CSD)		For example, Clearstream Banking AG, Frankfurt, acts as the officially recognized German bank for the central depository of securities under the Depotgesetz (the German Securities Deposit Act), among other things. In this function, it offers a wide range of post-trade services relating to securities issued in Germany and other countries, both as a CSD for securities eligible for collective safe custody and as a custodian for other securities[c]	An entity that (1) enables securities transactions to be processed and settled by book entry; (2) provides custodial services (e.g., the administration of corporate actions and redemptions); and (3) plays an active role in ensuring the integrity of securities issues. Securities can be held in a physical (but immobilized) form or in a dematerialized form (whereby they exist only as electronic records)[f]	

Clearing		The netting of receivables and liabilities arising from securities and derivatives transactions; determination of the bilateral net debt of buyers and sellers[c]	A system used to settle mutual indebtedness between a number of organizations (banks, brokers, etc.). All claims are set against one another, a balance is struck at agreed intervals and only the differences are settled. Used frequently in securities clearing. Securities clearing: Reciprocal settlement (clearing) of the securities transactions of banks or brokers via common clearing houses. This not only facilitates the technical processing of securities trades but also reduces book and account entries to a minimum. Futures clearing covers all securities bought and sold forward, whereas spot clearing covers spot transaction[e]	Clearing means the process of establishing positions, including the calculation of net obligations, and ensuring that financial instruments, cash, or both are available to secure the exposures arising from those positions. Regulation (EU) No 648/2012 Art. 2 (3)
		In the case of derivatives, the management of open derivatives positions including their netting. Termination of derivatives contracts is also part of derivatives clearing involving the establishment of final positions for settlement. Mitigating the counterparty risks on open derivatives positions is the most important aspect of derivatives clearing. As derivatives contracts can have long maturities, clearing plays a crucial role in the derivatives value chain and is considerably more complex than, for example, the clearing of cash equities[b]		
Clearing house	Organizations […] that are exchange affiliated and facilitate the validation, delivery, and settlement of securities transactions		Also clearing organization: Institution which settles the payment of financial transactions between members. Used to rationalize securities trading and bank payments. Local clearing houses exist in all major banking centers, often alongside other organizations working on a national or international basis[e]	
Clearing member		For example, a Eurex participant who satisfies the capital requirements of Eurex Clearing AG and who has applied for, and been granted, a licence to clear derivatives traded on Eurex Exchange. A DCM may clear its own transactions and those of its customers, as well as those of 100 % affiliated group companies that do not hold a clearing licence[c]		Clearing member means an undertaking which participates in a CCP and which is responsible for discharging the financial obligations arising from that participation. Regulation (EU) No 648/2012 Art. 2 (14)
		Market participant holding a clearing license at a CCP[b]		
Client		Clients of clearing members can access centrally cleared financial instruments via their clearing members[b]		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Client asset protection/ segregation		A CCP keeps separate records and accounts for the assets and positions of clearing members' clients. In case of a clearing member default, the clients' assets and positions are protected and can be transferred to another clearing member[b]		
Collateral/ collateral management		Collateral comprises assets given as a guarantee by a borrower (collateral provider) to secure a loan or other financial exposures and which are subject to utilization by the lender (collateral taker) in the event of default. Collateral management encompasses the administration and custody of deposited collateral to cover financial exposures, for example resulting from securities lending transactions or derivatives transactions[c]		
Collaterized debt obligation	Investment-grade bond backed by a pool of variously rated bonds, including Junk Bonds. CBOs are similar in concept to collateralized mortgage obligations but differ in that CBOs represent different degrees of credit quality rather than different maturity. Underwriters of CBOs package a large and diversified pool of bonds, including high-risk, high-yield junk bonds, which is then separated into tiers, called tranches […]			
Co-location		A service offered by trading system operators that allows exchange members to install front end infrastructure in close proximity to the exchange systems, e.g., in the same data center location, as a way to minimize latency in the data transmission between the member front end and the exchange back end		

Commercial bank money			Commercial bank liabilities that take the form of deposits held at a commercial bank which can be used for settlement purposes[f]	
Computer operating system		A software that manages computer hardware and software resources and provides common services for other computer programs. Microsoft Windows or Linux, for example, are operating systems. Higher order application programs usually require an operating system to function		
Contagion	The spreading of an economic crisis from one geographical area to another […]			
Continuous trading		The safekeeping and administration of securities for others. A custody account (similar to an account for money transactions) is established for each customer. The account information includes details of the types, nominal values or quantities, volumes, etc. of the securities held, as well as the name and address of the account holder[b]	Also regular trading. The third session of stock market trading, during which new orders are continually matched to existing ones in the order book according to the matching rules, and orders that cannot be executed are entered in the order book[e]	
Corporate action/ event			An action or event decided by the issuer of a security which has an impact on the holders of that security. This may be optional, in which case those holders have a choice (for example, they may have the right to purchase more shares, subject to conditions specified by the issuer). Alternatively, it may be mandatory, whereby those holders have no choice (e.g., in the case of a dividend payment or stock split). Corporate actions can relate to cash payments (e.g., dividends or bonuses) or the registration of rights (subscription rights, partial rights, splits, mergers, etc.)[f]	

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Counterparty		The opposite party to a financial transaction. Normally, the counterparty of the buyer of a contract is the seller of that contract. In the case of securities cleared by a CCP, the clearing house (e.g., Eurex Clearing AG) acts as the central counterparty to each party of a transaction, thereby removing counterparty risk from the members[c]		
Credit risk			The risk that a counterparty will not settle the full value of an obligation—neither when it becomes due, nor at any time thereafter. Credit risk includes replacement cost risk and principal risk. It also includes the risk of the settlement bank failing[f]	
Custodian	Bank or other financial institution that keeps custody of stock certificates and other assets of a mutual fund, individual or mutual client		Also: depositary; depository. Bank or broker where the valuables contained in a custody account are stored for safekeeping[e]	
Custody	Legal responsibility for someone else's assets. Term implies management as well as safekeeping	The safekeeping and administration of securities for others. A custody account (similar to an account for money transactions) is established for each customer. The account information includes details of the types, nominal values or quantities, volumes, etc. of the securities held, as well as the name and address of the account holder[c]		
Dark pool	The huge amount of liquidity created by institutional investors, mostly represented by block trades made off-board, the details of which are not available to the public			

Dealer	Individual or firm acting as principal in a securities transaction. Principals trade for their own account and risk. When buying from a broker acting as dealer, a customer receives securities from the firm's inventory; the confirmation must enclose this. When specialist trade for their own account, as they must as part of their responsibility for maintaining an orderly market, they act as dealers. Since most brokerage firms operate both as brokers and as principals, the term broker-dealer is commonly used	Persons or companies that function as intermediaries in the purchase and sale of securities. Dealers are also authorized to trade securities for their own account[a]	Someone who, in addition to acting as a buying and selling agent for others, buys and sells for his or her own account[e]	
Default	Failure of a debtor to make timely payments of interest and principal as they come due or to meet some other provision of a bond indenture. In the event of default, bondholders may make claims against the assets of the issuer in order to recoup their principal		Failure to make required debt payments on a timely basis or to comply with other conditions of an obligation or agreement[e]	
Delivery versus payment (DVP)	Securities industry procedure, common with institutional accounts, whereby delivery of securities sold is made to the buying customer's bank in exchange for payment, usually in the form of cash	The irrevocable exchange of securities (the delivery) and cash value (the payment) to settle a transaction. True delivery against payment involves the simultaneous exchange of securities and cash, as in the case of internal transactions between CBL counterparties[d]		
Dematerialized		Securities held on an institution's books in electronic form, without any registered certificates issued[d]		

(continued)

Term	Barron’s	Deutsche Börse group	Third party	Regulatory documents
Demutualization	Conversion of a member-owned institution or mutually owned company to another form of organization, usually a shareholder-owned company. Demutualization which is usually done to make access to capital easier has been a trend in the insurance industry and has been done in the case of several stock exchanges			
Disaster recovery		Disaster recovery involves a set of policies and procedures that enable the recovery or continued operation of vital technology infrastructure and systems following a natural disaster or other business disruption		
Dodd-Frank Wall Street Reform and Consumer Protection Act	Massive financial law passed in 2010 designed to prevent another financial crisis similar to what occurred in 2008. The bill, named after Connecticut Senator Christopher J. Dodd and Massachusetts Congressman Barney Frank, is designed to monitor and lower risk throughout the banking and financial system. It created new government agencies, including the Financial Stability Oversight Council, Office of Financial Research, and Orderly Liquidation Authority to monitor the financial strength of financial firms considered “too big to fail” and, if necessary, to step in to liquidate them in an orderly fashion to avoid taxpayer bailouts […]			Dodd-Frank is a legislative act in the USA that promotes “the financial stability of the United States by improving accountability and transparency in the financial system, to end ‘too big to fail’, to protect the American taxpayer by ending bailouts, to protect consumers from abusive financial services practices, and for other purposes.”—H. R. 4173 (111th)
Domestic security		A security issued in a national market[d]		

Eurobond	Bond denominated in US dollars or other currencies and sold to investors outside the country whose currency is used. The bonds are usually issued by large underwriting groups composed of banks and issuing houses from many countries	Bond issued by a borrower outside of a domestic market, denominated in a Eurocurrency, underwritten and sold by an international syndicate of financial institutions. The securities are generally listed, but not traded, on either the London or the Luxembourg Stock Exchange. Private Placements, which are not offered for public sale outside the initial group of allottees, are not usually listed. Eurobonds are not generally subject to the tax and other registration requirements of any national market, but may be subject to selling restrictions, particularly to US residents, before being seasoned[d]	Bond issued on the Euromarket in a currency other than that of the issuer's country of origin[e]	
European Market Infrastructure Regulation (EMIR)		EMIR regulates OTC derivatives, central counterparties (CCPs), and trade repositories (TR); it aims to improve security and integrity within the OTC derivatives market by promoting transparency and reducing risk. Among other things, the regulation wants to achieve this by introducing a clearing obligation for eligible OTC derivatives, measures to reduce counterparty credit risk and operational risk for OTC derivatives not cleared via CCPs, as well as disclosure requirements for all derivatives. It also establishes general requirements for CCPs and trade repositories[c]		
European single market				An area of free movement for goods, people, services and capital, the internal market has been further developed since 1993 by the consolidation of economic integration, the Euro and solidarity and cohesion policies.—COM(2011) 206
European Supervisory Authorities		Established in 2010 as part of the ESFS. Current ESAs are EBA, ESMA, and EIOPA responsible for micro-prudential oversight within the EU alongside the ESRB		—*Regulation* (*EU*) *No 1092/2010*

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
European System of Financial Supervision				Established in 2010 following the recommendation of the Larosière report to bring together actors of financial supervision at national and at Union level to act as a network. The network comprises the ESRB for macro-prudential oversight and the ESAs for micro-prudential oversight.—*Regulation* (*EU*) *No 1092/2010*
European Systemic Risk Board			The ESRB is part of the European System of Financial Supervision (ESFS), responsible for the macro-prudential oversight of the financial system within the Union in order to contribute to the prevention or mitigation of systemic risks to financial stability in the Union that arise from developments within the financial system and taking into account macroeconomic developments, so as to avoid periods of widespread financial distress (https://www.esrb.europa.eu/)	—Regulation (EU) No 1092/2010
Ex-dividend	Interval between the announcement and the payment of the next dividend. An investor who buys shares during that interval is not entitled to the dividend. Typically, a stock's price moves up by the dollar amount of the dividend the ex-dividend date approaches, then falls by the amount of the dividend after that date		Execution of a trade on an ex-dividend basis entitles the seller to the dividend whilst the buyer does not gain entitlement[h]	
Fairness		Fairness in the context of technical access to exchange systems means that members have equal access to a transparently structured menu of connectivity options		

Flash crash	An unexpected, precipitous drop in stock prices, often caused by a programming or other technical trading glitch. The first flash crash took place on May 6, 2010, when many stock prices plunged 10 % or more in a matter of minutes for no apparent reason and then quickly rebounded to their previous levels, causing many trades to be cancelled thereafter. The SEC subsequently issued new rules that automatically stop trading in any S&P 500 stock whose price moves by more than 10 % in a 5-min span			
Flow control		Flow control in electronic networks and trading systems is the management of data transmission between computers or devices or between nodes in a network so that the data can be handled at an efficient pace. Too much data arriving before a device can handle it causes data overflow, meaning the data is either lost or must be retransmitted		
Foreign bond	A debt security issued by a foreign entity in a domestic market in the domestic market's currency. Attraction to domestic investors is international diversification without currency exchange risk	Security issued by a borrower in another country's national market, usually denominated in the lending country's currency, and distributed or underwritten in the lending country through a management group of that country. Examples: Yankee, Samurai, Bulldog, and Matador bonds[d]		
Free float		The float refers to shares that are not owned by major shareholders, and can therefore be acquired and traded by the general public. As a rule, the larger the float, the easier it is for investors to buy and sell the stock[a]	The part of a company's shares which is not held as a long-term investment by the founder, management, etc. and can therefore be traded on a stock exchange at any time, making it available to the investing public[e]	
Free of payment (FOP)		A delivery of securities which is not linked to a corresponding transfer of funds, unlike DVP[c]		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Front end		A front-end application is a software interface with which system users may interact directly as opposed to a back end, which may only be accessed via other intermediary systems		
Front office	Sales personnel in a brokerage, insurance, or other financial services operation. Front-office workers produce revenue, in contrast to back-office workers, who perform administrative and other support functions for the front office	In a trading environment, the term front office describes the trading desk personnel who interact with clients, make trading decisions, enter orders, etc. The front office is strictly separated from the back office which carries out administration, support, and control activities		
Fungibility		Interchangeability of currencies, rights, and goods, e.g., securities[a]	A characteristic of securities which are substitutable on account of their being identical[f]	
General clearing member		A Eurex participant who satisfies the capital requirements of Eurex Clearing AG and who has applied for, and been granted, a licence to clear derivatives traded on Eurex. AGCM may clear its own transactions, those of its customers, and those of market participants who do not hold a clearing licence (non-clearing members)[c]		

Glass-Steagall-Act	Legislation passed by Congress authorizing deposit insurance and prohibiting commercial banks from owning full-service brokerage firms. Under Glass-Steagall, these banks were prohibited from investment banking activities, such as underwriting corporate securities or municipal revenue bonds. The law was designed to insulate bank depositors from the risk involved when a bank deals in securities and to prevent a bank collapse like the one that occurred during the Great Depression. The original separation of commercial and investment banking had already significantly eroded when, on November, 12 1999 the Financial Services Modernization Act was signed into law, repealing parts of the 1933 Glass-Steagall Act and the 1956 Bank Holding Company Act and effectively allowing banks, brokers, and insurers into each other's businesses. The 1999 Act allows banks to affiliate with securities firms and insurers through a holding company structure and permits nationally chartered banks to engage in most financial activities through direct subsidiaries. While provisions of Glass-Steagall continue to restrict banks from most underwriting activities and securities firms from taking deposits, these restrictions apply only to the banks and securities firms, not to their Financial Holding Company affiliates and are, therefore, technical	Also known as Banking Act of 1933—an act under US law published on 16 June 1933 "to provide for the safer and more effective use of the assets of banks, to regulate interbank control, to prevent the undue diversion of funds into speculative operations, and for other purposes." The Act was especially prominent for including a separation of commercial and investment banking (c.f. sections, 16, 20, 21 and 32). This specification was repealed with the passage of the Gramm–Leach–Bliley Act (GLBA), allowing the consolidation of commercial banks, investment banks, securities firms, and insurance companies		H.R. 5661 (73rd)

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Haircut			A risk control measure applied to underlying assets whereby the value of those underlying assets is calculated as the market value of the assets reduced by a certain percentage (the "haircut"). Haircuts are applied by a collateral taker in order to protect itself from losses resulting from declines in the market value of a security in the event that it needs to liquidate that collateral[f]	
Hardware		The machines, wiring, and other physical components of a computer or other electronic system		
High-frequency trading (HFT)/ algorithmic trading	Algorithmic trading, also called "flash" trading, done in microseconds (one millionth of a second) using supercomputers and making it possible for traders to exploit subtle and obscure market signals and inefficiencies and to earn rebates paid by exchanges for quickly providing shares when needed. Regulators are concerned that unscrupulous operators could trade ahead of orders and affect prices. High-frequency trading accounts for a huge amount of the volume of trading on most stock exchanges	Form of algorithmic trading. During high-frequency trading, orders for the purchase and sale of financial instruments are determined through the use of high-frequency algorithmic trading strategies and are rapidly transmitted to a trading venue. High-frequency traders trade for their own account and place a high number of intraday orders, quotes, etc.[c]		Algorithmic trading means trading in financial instruments where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price, or quantity of the order or how to manage the order after its submission, with limited or no human intervention, and does not include any system that is only used for the purpose of routing orders to one or more trading venues or for the processing of orders involving no determination of any trading parameters or for the confirmation of orders or the post-trade processing of executed transactions. Directive 2014/65 EU Art. 4 (39)
Immobilization		The collective storage of securities in a vault in order to eliminate physical movement of Certificates or documents of ownership when transfer of ownership occurs[d]		

Initial public offering	Corporation's first offering of stock to the public. IPOs are almost invariably an opportunity for the existing investors and participating venture capitalists to make big profits, since for the first time their shares will be given a market value reflecting expectations for the company's future growth	A type of public offering in which shares of stock in a company are sold to investors. Through this process a private company transforms into a public company. After the IPO, when shares trade freely in the open market, money passes between public investors in the secondary markets	Also known as IPO or going public. First sale to the general public, by issuing and listing shares, of stock in a company whose equity had previously been held by a restricted group of persons[e]	
Interest rate swaps (IRS)			Swap transaction whereby the two parties undertake to swap variable interest payments for fixed interest payments in respect of a certain nominal value at regular intervals over a specified period[e]	
Intermediary	Person or institution empowered to make investment decisions for others. Some examples are banks, savings, and loan institutions, insurance companies, brokerage firms, mutual funds, and credit unions. These specialists are knowledgeable about investment alternatives and can achieve a higher return than the average investor can		Natural person or legal entity that accepts assets from other persons keeps them in custody or helps to invest or assign them. A fiduciary is a financial intermediary, for instance, as are banks and brokers in the case of stock exchange transactions[e]	
International central securities depository (ICSD)		A central securities depository that provides clearance and settlement of transactions in global and international securities and domestic securities traded across borders[d]		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
International Securities Identification Number (ISIN)		International Securities Identification Number. A coding system developed by the ISO with the purpose of creating one unique number on a worldwide basis for identifying securities in accordance with ISO standard 6166. The ISIN for each security consists of a 12-digit alphanumeric code. The prefix is a two-letter country code (or XS in the case of Euro-instruments, for which CBL and Euroclear act as numbering agents). The basic number is a nine-digit alphanumeric code, which is the common code of CBL and Euroclear in the case of Euro-instruments. The final digit is a numeric check digit computed from the preceding digits[d]	A 12-digit number consisting of three elements used to identify securities. The first two places are reserved for a combination of letters indicating the country of origin (e.g., DE for Germany, LU for Luxembourg). These are followed by a national identification number consisting of up to ten digits[e]	
International security		A security issued and distributed simultaneously through an international management group in one or more domestic or international markets[d]		
Investor	Party who puts money at risk; may be an individual or an institutional investor		Person or institution purchasing securities or tangible assets for the purpose of earning a return or receiving the monetary value[e]	
Issuer	Legal entity that has the power to issue and distribute a security. Issuers include corporations, municipalities, foreign and domestic governments and their agencies, and investment trusts. Issuers of stock are responsible for reporting on corporate developments to shareholders and paying dividends once declared. Issuers of bonds are committed to making timely payments of interest and principal to bondholders	A company, or government body, that borrows or raises funds through the sale of securities[d]	Also: borrower. Private sector corporation or public authority raising capital for its own use by issuing securities on the public market[e]	
Latency		The time between entering and processing an order. Electronic trading platforms that execute orders quickly with a short response and processing time have a low latency[a]		

Lead manager		The financial institution that is primarily responsible for the overall coordination, distribution, and documentation of a primary market (new) issue. The lead manager is primarily responsible to the borrower or issuer for selecting the co-managers, determining the terms of the issue, and selecting underwriters and the members of the selling group[d]	Also: lead bank. Bank acting as the leader of an issuing or lending syndicate, managing the transaction. It is normally responsible for contact with the borrower, for the structure, organization, and text of the loan agreement and prospectus, plus the composition of the issuing syndicate and the selling group and, ultimately, for the placement itself[e]	
Legal risk			The risk of a loss being incurred on account of the unexpected application of a law or regulation, or because a contract cannot be enforced[f]	
Limit order	Order to buy or sell a security or commodity at a specific price or better. The broker will execute the trade only within the price restriction			
Lines of defence (LoD)		The multiple risk mitigation layers of a CCP, often referred to as the CCP risk waterfall. The lines of defence of CCPs are prudently scaled to meet severe stress scenarios and ensure confidence that CCPs can guarantee contracts[b]		
Liquidity		The liquidity of a security is a function of the number of shares or units in circulation and the number of market participants who are willing to buy or sell them. If a security is liquid, this means that both supply and demand are enough to ensure that a trade—another term for the simultaneous purchase and sale of a security—can take place at any time[a]		
Liquidity risk	The risk of being unable to sell an asset quickly at its fair market value. Assets with active markets, such as listed stocks, have lower liquidity risk than assets with fewer potential buyers, such as paintings	Liquidity risk can mean either (A) market or (B) funding liquidity risk. (A) Market liquidity risk materialises when financial assets cannot be sold rapidly at the presumed market value. (B) Funding liquidity risk is the risk that a counterparty has insufficient funds to meet its financial obligations when they are due[b]	The risk that a counterparty will not settle an obligation in full when it becomes due. Liquidity risk does not imply that a counterparty or participant is insolvent, since it may be able to effect the required settlement at some unspecified time thereafter[f]	

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
listed stock	Security, stock, or bond that has been accepted for trading by one of the organized and registered securities exchanges […]. Generally, the advantages of being listed are that the exchanges provide (1) an orderly marketplace; (2) liquidity; (3) fair price determination; (4) accurate and continuous reporting; (5) information on listed companies; and (6) strict regulations for the protection of securities holders			
Long position	Ownership of a security, giving the investor the right to transfer ownership to someone else by sale or by gift; the right to receive any income paid by the security; and the right to any profits or losses as the securitie's value changes			
Macro-prudential supervision		The monitoring of the macroeconomic environment in order to identify, assess, and prevent systemic risks to financial stability. In the EU this form of supervision is provided by the ESRB		
Margin call			Demand for additional funds to increase the value of the cover for loans collateralized by securities or for financial futures whose prices have fallen below a predetermined level[e]	
Margin collateral		Used to secure a transaction. In the derivatives market, collateralization plays an important role to manage counterparty risk[b]		

Term	Barron’s	Deutsche Börse group	Third party	Regulatory documents
Market making/ market maker	Make a market: Maintain firm bid and offer prices in a given security by standing ready to buy or sell round lots at publicly quoted prices. The dealer is called a market maker in the over-the-counter market and a designated market maker (DMM) on the exchanges	Market makers guarantee that a security can be traded on an ongoing basis, thereby ensuring the liquidity and viability of the market. Through their readiness to step in as a counterparty at any time, market makers compensate for the inconsistent order flows of investors and stabilize the market when temporary imbalances occur. They are usually assigned to support trading in inactive stocks[a]		Market maker means a person who holds himself out on the financial markets on a continuous basis as being willing to deal on own account by buying and selling financial instruments against that person’s proprietary capital at prices defined by that person. Directive 2014/65 EU Art. 4 (7)
Market order	Order to buy or sell a security at the best available price. Most orders on the exchanges are market orders			
Matching		Automatic offsetting and execution of orders in an electronic trading system[a]		
Matching engine		The functionality of a trading back end that manages order books and matches orders for execution according to the matching rules of the particular market model		
Microburst		A sudden and short-lived surge in the market data volume that must be handled by exchange systems		
Micro-prudential supervision		The specific supervision of financial institutions and individual markets by the ESAs in close cooperation with national authorities, which remain in charge of regular, ongoing supervisory practices. Three supranational organizations oversee micro-prudential supervision within the EU: EBA, EIOPA, and ESMA working together in the Joint Committee of the European Supervisory Authorities (Joint Committee) provided for by Article 54 of Regulation (EU) No. 1093/2010, of Regulation (EU) No. 1094/2010 and of Regulation (EU) No. 1095/2010		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Microwave		An electromagnetic wave with a wavelength in the range 0.001–0.3 m, shorter than that of a normal radio wave but longer than those of infrared radiation. Microwaves are used in radar, in data communications, and for many other commercial and industrial processes		
Middle office		In a trading environment, the middle office is interposed between the trading activity in the front office and the administrative functions of the back office. Its main mission is the real-time enforcement of limits and other risk management duties		
Mortgage-backed security	Security backed by mortgages. […] Investors receive payments out of the interest and principal on the underlying mortgages […]			
Multilateral trading facility (MTF)		Securities firm or market operator that represents the interests of a large number of persons in the buying and selling of financial instruments within the system. It applies defined provisions so as to lead to an agreement on the purchase of these financial instruments[c]	A multilateral system which brings together multiple third-party buying and selling of interests in financial instruments—in the system and in accordance with nondiscretionary rules—in a way that results in a contract in accordance with the provisions of MiFID[e]	[…] means a multilateral system, operated by an investment firm or a market operator, which brings together multiple third-party buying and selling interests in financial instruments—in the system and in accordance with nondiscretionary rules—in a way that results in a contract […]. Directive 2014/65 EU Art 4 (22)
Mutualization		The distribution of losses across the parties active in a market segment. CCP members agree to mutualize losses among themselves should the losses exceed the collateral provided by the defaulting member and the CCPs own contributions. This is a stabilising factor equivalent to an insurance scheme since a large amount of security is available at a marginal cost to the affected non-defaulting members[b]		

Netting		Offsetting buy and sell positions over a given period of time so that market participants only have to settle the balance. If two parties agree to net their positions, this is called bilateral netting. Central counterparties even allow the netting of three or more parties' positions, which is called multilateral netting[b]	Mutual offsetting of claims and liabilities from identical types of transaction between two parties. Important in connection with the capital adequacy requirements applicable to banks. In industry: a method of optimizing intra-group capital transfers via a central account of net credits[e]	
			In the context of clearing or settlement systems, the agreed offsetting of mutual obligations by participants in a system. This process involves the calculation of net settlement positions and their legal reduction to a (bilateral or multilateral) net amount[f]	
Network Time Protocol		The Network Time Protocol is the standard time protocol used to synchronize the clocks of the single devices in a complex electronic network with a precision in the range of milliseconds		
New issue	Stock or bond being offered to the public for the first time, the distribution of which is covered by Securities and Exchange Commission (SEC) rules. New issues may be initial public offerings by previously private companies or additional stock or bond issues by companies already public and often listed on the exchanges […]	A security offered for the first time[d]		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Omnibus account		Omnibus account holds securities on behalf of more than one beneficial owner and is managed by a single operating institution[c]	A single account for the commingled funds or positions of multiple parties. A clearing member will often maintain an omnibus account at the clearing house for all of the clearing member's clients. In this case, the clearing member is responsible for maintaining account records for individual clients[g]	
Operational risk		Temporary or permanent disruption of a market participant's or a CCP's operations. Central aspects of the CCP's operations are its IT infrastructure, facilities, and workforce[b]	The risk of human error or a breakdown of some component of the hardware, software, or communications systems that are crucial to settlement[g]	
Order book		An order book is used to pool, compare, and match the volumes and prices of buy and sell orders for a particular security. Thus, in auction-based trading, the order book supports the price determination procedure. Market orders, which are to be executed at the best available price, are given first priority in the order book. The remaining orders are sorted and listed, with the bid prices in ascending order and the ask prices in descending order[a]		
Organized trading facilities (OTF)		New type of trading venue which is set to be introduced within the context of the review of the financial markets directive MiFID, in addition to the existing categories of the regulated market and of multilateral trading facilities[c]		[…] means a multilateral system which is not a regulated market or an MTF and in which multiple third parties are buying and selling interests in bonds, structured finance products, emission allowances, or derivatives are able to interact in the system in a way that results in a contract […]. Directive 2014/65 EU Art 4 (23)

Over the counter (OTC)	Market in which securities transactions are conducted through a telephone or computer network connecting dealers in stocks and bonds, rather than on the floor of an exchange	Also: off-exchange. Describes transactions between two or more trading parties that are not conducted on a regulated market. The OTC segment accounts for by far the largest part of the derivatives market[c]	Also: off exchange. Securities transactions outside the stock market on the so-called over-the-counter market[e] A method of trading that does not involve a regulated market. In over-the-counter markets, participants trade directly with each other, typically through telephone or computer links[f]	Execution [of trades] […] which does not take place on a regulated market as within the meaning of Article 4(1) (14) of Directive 2004/39/EC or on a third-country market considered as equivalent to a regulated market in accordance with Article 19(6) of Directive 2004/39/EC. Regulation (EU) No 648/2012 Art. 2 (14)
Paying agent	Agent, usually a bank that receives funds from an issuer of bonds or stock and in turn pays principal and interest to bondholders and dividends to stockholders, usually charging a fee for the service. Sometimes called distributing agent	The financial institution(s) responsible for the task of making due payments of principal and interest to the holders of an issue of a security against presentation of the security or its coupons. The principal paying agent is responsible for collecting the money due from the Issuer and for coordinating the distribution of payments, through the sub-paying agents, to the holders of the issue, on demand[d]	Agent or bank appointed by the issuer of securities to carry out all ongoing transactions that arise for the owners of the said securities, e.g., payment of dividends and interest coupons[e]	
Payment system (PS)			A payment system consists of a set of instruments, banking procedures, and, typically, interbank fund transfer systems that ensure the circulation of money[g]	
Performance bond	Surety bond given by one party to another, protecting the second party against loss in the event the terms of a contract are not fulfilled. The surety company is primarily liable with the principal (the contractor) for non-performance […]		Bond, usually issued by a bank, guaranteeing specific monetary payment to a beneficiary if the purchaser or maker fails to perform or acts in violation of a contract. In the USA, a performance bond puts the issuer under obligation to render the performance himself[e]	

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Pledging	Transferring property, such as securities or the cash surrender value of a life insurance, to a lender or creditor as collateral for an obligation. Pledge and hypothecate are synonymous as they do not involve transfer of title			
Pool factor		Factor to be applied to the nominal amount of a security, reflecting partial redemptions, to obtain the issue outstanding amount[d]		
Post-trade		A summary term for back office processes that intervene after a trade has been agreed or an order execution in an electronic system has occurred. Post-trade functions include matching trade confirmations, issuing settlement instructions, resolving discrepancies, etc.		
Predictability		Consistent system performance that is independent of current transmission and processing load. Predictability under peak loads requires large spare capacity compared to average transmission and processing volumes		
Price discovery	Free market process by which an illiquid asset eventually attracts a buyer and discovers what price it will fetch			
Primary market	Market for new issues of securities, as distinguished from the secondary market, where previously issued securities are bought and sold. A market is primary if the proceeds of sales go to the issuer of the securities sold. The term also applies to government securities auctions and to opening option and futures contract sales	A market in which new equity or debt securities are issued by companies or governments to obtain financing. Primary markets are facilitated by underwriting groups, which consist of investment banks that will set a beginning price range for a given security and then oversee its sale directly to investors	The market on which new securities are launched and enter into circulation[e]	

Queue		Short-term buffer for storing data while awaiting processing		
Random walk	Theory about the movement of stock and commodity futures prices hypothesizing that past prices are of no use in forecasting future price movements. According to the theory, stock prices reflect reactions to information coming to the market in random fashion so they are no more predicatble than the walking pattern of a drunken person […]			
Real time		Relates to a system where the actual time elapsed in the performance of a computation is negligibly small so that the result is available virtually immediately for further processing		
Real-time gross settlement (RTGS)			A settlement system in which processing and settlement take place on a transaction-by-transaction basis in real time[e]	
Record date		The date, established by an issuer of a security, used by CBL to determine, at the end of that day (that is, after end of day processing) the holders that are entitled to a corporate action[d]		
Redemption	Repayment of a debt security or preferred stock issue, at or before maturity, at par or at premium price	In accordance with the terms and conditions of an issue, partial or full return of the debt certificates of an issuer, for cancellation of the certificates against payment to the holder[d]		
Redundancy		Relates to the installation of system and network components in duplicates or higher multiples that may instantly replace each other, so as to avoid any single point of failure		

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Registered security	Security whose owner's name is recorded on the books of the issuer or the issuer's agents, called a registrar—for example a registered bond as opposed to a bearer bond, the former being transferrable only by endorsement, the latter payable by the holder		A security where the issuer (or a registrar acting on the issuer's behalf) maintains a record of owners of the security; this usually requires that when securities are sold, the seller's name is replaced by the buyer's name on the register. Typically, equities rather than bonds are issued in registered form. See Bearer form[h]	
Regulated market				"[…] means a multilateral system operated and/or managed by a market operator, which brings together or facilitates the bringing together of multiple third-party buying and selling interests in financial instruments—in the system and in accordance with its non-discretionary rules—in a way that results in a contract, in respect of the financial instruments admitted to trading under its rules and/or systems, and which is authorised and functions regularly […]"—*Directive 2014/65 EU Art 4 (21)*
Regulation	The process of rulemaking and legislation underlying a supervisory framework			
Reliability		The probability that a system will produce correct outputs. Reliability is enhanced by features that help to avoid, detect, and repair faults		

Repurchase agreement (Repo)	Agreement between a seller and buyer […], whereby the seller agrees to purchase the securities at an agreed upon price and, usually, at a stated time. Repos, also called RPs or buybacks, are widely used both as a money market investment vehicle and as an instrument of Federal Reserve monetary policy. Where a repurchase agreement is used as a short-term investment, a government securities dealer, usually a bank, borrows from an investor, typically a corporation with excess cash, to finance its inventory, using the securities as collateral. Such RPs may have a fixed maturity date or be open repos, callable at any time	A contract to execute two simultaneous transactions for the against payment sale of securities and their repurchase at a future date. Repo agreements are normally executed (by securities dealers) in lieu of borrowing funds against the delivery of the securities (which form part of the dealer's inventory or trading position and which act as collateral to the lender). The terms of the agreement normally allow the seller to retain the rights of the holder of the security for coupon payments etc. The price at which the securities are repurchased and the period between sale and repurchase reflect money market terms and rates of interest on the loan[d]		
Safekeeping			The holding of physical securities on behalf of other parties[f]	
Screen-based trading		A trading style on an electronic trading system requiring direct human intervention for entering orders and authorizing trades		
Secondary market	Exchanges and over-the-counter markets where securities are bought and sold subsequent to original issuance, which took place in the primary market. Proceeds of secondary market sales accrue to the selling dealers and investors, not to the companies that originate the securities	In the secondary market, securities are sold by and transferred from one investor to another. It is therefore important that the secondary market be highly liquid and transparent. The eligibility of stocks and bonds for trading in the secondary market is regulated by the Stock Exchange Act and the stock exchange rules and regulations. Foreign financial instruments can become eligible through the issue of depositary receipts. The derivatives market consists only of a secondary market for standardized derivative instruments (i.e., instruments with comparable contract specifications)[a]	Market on which the trading of securities takes place[e]	

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Securities		A financial instrument that represents: an ownership position in a publicly-traded company (equity) or a creditor relationship with governmental body or a company (bond). Securities are fungible, negotiable financial instruments that represent financial value	Instruments that signify an ownership position in a corporation (stocks or shares), a creditor relationship with a corporation or government body (bonds), or rights to ownership such as those embodied in options or subscription rights[e]	
Securities lending			The practice of a share or bond holder loaning its securities to a borrower for a fee, in order to enhance the lender's return on its investment[h]	
Securities settlement system (SSS)			A system which allows the transfer of securities, either free of payment (FOP) or against payment (delivery versus payment)[f]	
Securitization	Process of distributing risk by aggregating debt instruments in a pool, then issuing new securities backed by the pool			
Self-regulatory organization (SRO)	Principal means template by the federal securities laws for the enforcement of fair, ethical, and efficient practices in the securities and commodities futures industries			
Sell side		The part of the financial industry involved with the creation, promotion, analysis and sale of securities. Sell-side individuals and firms work to create and service stock products that will be made available to the buy side of the financial industry		
Server		A running instance of an application (software) capable of accepting requests from "client" programs and giving responses accordingly. Multiple servers can run on a single computer, but for security reasons, a server application is typically installed on dedicated computer hardware, which may then also be referred to as "the server"		

Settlement	In general, a resolution of differences among various parties. […] Securities: Conclusion of a securities transaction in which a broker/dealer pays for securities bought for a customer or delivers securities sold and receives payment from the buyer's broker. […] Futures/options: the final price, established by exchange rule, for the prices prevailing during the closing period and upon which futures contracts are marked to market	The completion of an exchange transaction, i.e., the transfer of money and traded securities from the seller to the buyer and vice versa. Within Deutsche Börse Group, Clearstream is responsible for this post-trading function[c]		
		In the case of derivatives, the sole payment of cash to fulfil the obligation arising from a derivatives contract (cash settlement) or the payment of cash for an underlying and the delivery of the underlying in return (physical delivery)[b]		
Shadow banking system	The miscellany of non-deposit taking financial institutions that exist between investors and borrowers but are not subject to the reserve requirements and other disciplines that banks are […]			
Short position	Stock shares that an individual has sold short (by delivery of borrowed certificates) and has not covered as as of a particular date			
Society for Worldwide Interbank Financial Telecommunication (SWIFT)			A cooperative organization created and owned by banks that operates a network which facilitates the exchange of payment and other financial messages between financial institutions (including broker-dealers and securities companies) throughout the world. A SWIFT payment message is an instruction to transfer funds; the exchange of funds (settlement) subsequently takes place over a payment system or through correspondent banking relationships[g]	

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Software		The programs and instructions used to run a computer system. Generally distinguished from the hardware, i.e., its physical components		
Stock exchange	Organized marketplace in which stocks, common stock equivalents, and bonds are traded by members of the exchange, acting both as agents (brokers) and as principals (dealers or traders). Most exchanges have a physical location where brokers and dealers meet to execute orders from institutional and individual investors to buy and sell securities. Each exchange has its own requirements for membership […]	Organized market for securities trading. Exchange trading takes place at established times, with the exchange itself performing the following main functions: • Bringing together • Supply and demand (market function) • Creating an environment in which companies can raise capital by issuing securities (mobilization function) • Guaranteeing that securities can be sold and transferred at any time (substitution function) • Determining the current market price for an individual stock, and thus the market value of the company in question (valuation function)[a]		
Straight-through processing (STP)	Direct exchange of cash for securities, common with cross-border transactions where settlement is often costly	The swift, safe and efficient processing of a securities transaction, from order placement to delivery versus payment and to the subsequent safe custody of the securities[c]	The automated end-to-end processing of trades/payment transfers—including, where relevant, the automated completion of confirmation, matching, generation, clearing and settlement of orders[f]	
Sub-custodian			Where one custodian (e.g., a global custodian) holds its securities through another custodian (e.g., a local custodian), the latter is known as a subcustodian[g]	
Supervision		The monitoring of the behavior of financial market participants and the enforcement of legislation		
Syndication account		An account opened in the name of a lead manager for the specific purpose of distributing the initial allotments and collecting the subscription proceeds of new issues[d]		

Systemic internalizer				"[…] means an investment firm which, on an organized, frequent systematic and substantial basis, deals on own account when executing client orders outside a regulated market, an MTF or an OTF without operating a multilateral system; […]. The definition of a systematic internalizer shall apply only where the pre-set limits for a frequent and systematic basis and for a substantial basis are both crossed or where an investment firm chooses to opt-in under the systematic internalizer regime."—Directive 2014/65 EU Art 4 (20)
Systemic risk	The risk that a sudden shock in an interlinked and interdependent financial industry will lead to a market seizure or a domino effect of failure that puts the entire financial system at risk. There are two types of systemic risk: (1) the risk of sudden, near-term systemic seizures or cascading failures, and (2) the longer term risk that our system will unintentionally favor large systemically important institutions over smaller, more nimble competitors, reducing the system's ability to innovate and adapt to change	The risk that the failure of one market participant has adverse effects on other market participants, destabilizing the market as a whole[b]	The risk that the inability of one participant to meet its obligations in a system will cause other participants to be unable to meet their obligations when they become due, potentially with spillover effects (e.g., significant liquidity or credit problems) threatening the stability of or confidence in the financial system. That inability to meet obligations can be caused by operational or financial problems[f]	A risk of disruption to financial services that is caused by an impairment of all or parts of the financial system and has the potential to have serious negative consequences for the real economy. (BIS, IMF, FSB definition: http://www.bis.org/publ/othp07.pdf) Risk of disruption in the financial system with the potential to have serious negative consequences for the internal market and the real economy. (EU definition: Regulation (EU) No 1092/2010)

(continued)

Term	Barron's	Deutsche Börse group	Third party	Regulatory documents
Systemically important institution			An institution is considered systemically if its malfunction causes widespread distress, either as a direct impact or as a trigger of broad contagion[g]	
TARGET2		Initiative to create a single platform for transmitting securities within the euro zone. The objective of this platform is to reduce the cost of cross-border securities settlement within the euro zone. It will be operated by the European Central Bank. "TARGET" is short for "Trans-European Automated Real-time Gross Settlement Express Transfer System"[c]	The real-time gross settlement system for the euro. TARGET2 settles payments in euro in central bank money and functions on the basis of a single IT platform, to which all payment orders are submitted for processing. This means that all payments are received in the same technical form. TARGET2 is legally structured as a multiplicity of RTGS systems (TARGET2 component systems)[f]	
TARGET2-Securities (T2S)			The Eurosystem's single technical platform enabling central securities depositories (CSDs) and national central banks to provide core, borderless, and neutral securities settlement services in central bank money in Europe[f]	
Time stamping		The process and methods of assigning a time stamp in a computer system. A time stamp is a record of the time of an event that is determined by way of a highly precise time reference and time dissemination protocol		
Too big to fail	Said of an organization whose failure would pose systemic risk			
Trade Repository				"[…] means a legal person that centrally collects and maintains the records of derivatives."—Regulation (EU) No 648/2012

Transaction	Execution of an order to buy or sell a security or commodity futures contract. After the buyer and seller have agreed on a price, the seller is obligated to deliver the security or commodity involved, and the buyer is obligated to accept it			
Transparency		Transparency in exchange trading systems relates to the disclosure of the order book, market data, audit trails, time stamps, and other system statistics that will enable members to fully analyze the performance of their own trading infrastructure and its interaction with the exchange back end		
Triparty repo			Repurchase agreement in which a third party (e.g., a custodian bank, a clearing house or a central securities depository (CSD)) is responsible for the management of the collateral during the life of the transaction[f]	
Two-sided market	Market in which both the bid and asked sides are firm, such as that which a specialist and others who make a market are required to maintain. Both buyers and sellers are thus assured of their ability to complete transactions			
Virtualization		In computing, virtualization refers to the act of creating a virtual (rather than actual) version of something. Software virtualization is used to simplify installation procedures and create encapsulated, self-contained software installations. An application can thus be deployed in a single file that is highly portable and platform independent. Virtualization comes at the expense of certain overheads and inefficiencies, which may be undesirable for the extreme performance requirements of electronic trading systems		

(continued)

Term	Barron’s	Deutsche Börse group	Third party	Regulatory documents
Volatility interruption		A safeguard mechanism in electronic exchange trading that improves price continuity, also called “circuit breaker.” For example, a volatility interruption will be initiated whenever the potential execution price of an order lies outside the dynamic and/or static price range around a reference price. A volatility interruption is typically resolved within a few minutes by soliciting additional orders and quotes from market participants and holding an intraday auction		
Waterfall		The multiple risk mitigation layers of a CCP[b]		
Zero footprint		A form of technical connection to an exchange back end that requires no special exchange software or hardware to be installed and maintained at the member site		

Note on sources:[a]Deutsche Börse glossary[b]Eurex glossary[c]Clearstream glossary[d]Clearstream customer handbook[e]UBS dictionary of banking[f]ECB glossary[g]BIS glossary[h]Michael Simmons: Securities Operations, John Wiley & Sons Ltd., 2002