John B. Guerard, Jr. *Editor*

Portfolio Construction, Measurement, and Efficiency

Essays in Honor of Jack Treynor

Springer

Portfolio Construction, Measurement, and Efficiency

John B. Guerard, Jr.
Editor

Portfolio Construction, Measurement, and Efficiency

Essays in Honor of Jack Treynor

Springer

Editor
John B. Guerard, Jr.
McKinley Capital Management
Anchorage, Alaska, USA

ISBN 978-3-319-33974-0 ISBN 978-3-319-33976-4 (eBook)
DOI 10.1007/978-3-319-33976-4

Library of Congress Control Number: 2016948853

© Springer International Publishing Switzerland 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Foreword

My memory, possibly faulty, is that I first met Jack Treynor at a meeting of the Institute for Quantitative Research in Finance ("Q Group"). It was most likely the meeting held at the Camelback Inn in Scottsdale, Arizona, in 1975. As I recall it was billed as a meeting to promote the reinvigoration of the group. The new Q Group leaders felt that there was a bigger role than previously for the group of like-minded practitioners and academics to collectively discuss how the new theories then emerging from academia could best be applied in practice. Many leading practitioners and several well-known academics were part of this new orientation. Among them was Jack.

Later, when I was fortunate to know him better and particularly after I got to see him in action presenting his ideas and insights on a broad range of topics from investor behavior and market structure to capital market theory I should not have been surprised to observe Jack deeply involved in this group. Not only was he then editor of the *Financial Analysts Journal (FAJ)*, but it was clear that he was held in very high esteem for his creativity and innovation by the attendees. It was obvious that he felt passionately about the theoretical foundations of financial economics—and indeed had played a major role in many of the most significant developments—but he was also equally passionate about the "best practices" of applying these theoretical developments in the real world. As a result his institutional influence was deep and profound, while his counsel and individual guidance widely sought.

As a young Ph.D. student in Operations Research, I was energized to observe Jack's effortless move between theory and practice, his powerful intellect and relentless drive to understand the intricacies of markets. He had studied mathematics at Haverford College and earned an M.B.A. with Distinction from Harvard Business School before being hired by the Operations Research Department at Arthur D. Little. There, as reported by Franco Modigliani, Jack had developed a capital asset pricing model (CAPM) in 1962 when searching for the applicable discount rate to be used in capital budgeting projects.

At the Q Group meeting where I first met Jack, Barr Rosenberg and I had been invited to give a talk on some of our work at Barra, Inc. The subject matter included some of the real-world applications of many of the theoretical innovations that Barr had made as a member of the Business School faculty at UC Berkeley, particularly his work in the estimation of investment risk using a multiple factor methodology.[1] This approach enabled the various components of risk to be decomposed and potentially separately measured and controlled. In particular, the approach enabled the contemporaneous prediction of risk that could be attributed to active management of the portfolio. At the time of the talk, we had built a portfolio analysis tool that estimated active risk and, by implication, suggested the efficacy of active management, as well as a portfolio optimization tool that built optimal equity portfolios with, for example, different risk tolerances towards the separate components of risk. This enabled portfolios to be constructed that exhibited, for example, high aversion towards specific risk but less aversion towards risk arising from style exposures (i.e., extra-market components of covariance) that were perceived to be associated with superior returns.

Jack, unsurprisingly, had also been involved in this area. In 1973 he had published an article entitled “How to Use Security Analysis to Improve Portfolio Selection,” coauthored with Fischer Black and published in the *Journal of Business.* The goal of the paper, as described, was to explore the link between the judgmental work of the security analyst and the more objective and quantitative approach of Markowitz and others. It was a theoretical paper that was based on a model that split an asset’s excess return into two components, a systematic component and an independent component of return. The authors further interpreted the independent return as the sum of an appraisal premium and residual error, which was their approach to modeling the value of security analysis. While they didn’t relate the independent return to microeconomic characteristics of companies or residual market factors as we attempted, their approach was also quite consistent and confirmatory of some of our thinking. In particular, our approach of measuring the active portfolio risk directly (as distinct from the total portfolio risk measurement) in order to understand the degree and effectiveness of active portfolio management.

This path of research led directly to the emergence at Barra, Inc. of new and more powerful investment technology based on the paradigm of a multiple factor model to understand and measure a wide variety of portfolio and investment strategy attributes. These included the measurement of portfolio risk, construction of tailored portfolios that were optimized to trade off sources of expected return versus the resulting active risk, identification and estimation of factors of expected return, the measurement of efficient measures of portfolio and factor-related performance, and the analysis of problems inherent in employing multiple managers for a single

[1] E.g., Rosenberg, B. (1974, March). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis,* 9(2), 263–274.

investment portfolio. These problems were thoroughly researched for several asset classes in multiple countries around the world.[2]

The widespread adoption of multiple factor models beginning in the 1980s sparked some confusion as to the relationship between factor models and risk estimation that results from them, and the CAPM. Jack weighed in on this debate in an important article published in the May-June 1993 issue of the *FAJ.* The article was entitled "In Defense of the CAPM" where he compared the Arbitrage Pricing Theory (APT), the CAPM in the context of a multiple factor structure. He writes provocatively that "there is nothing about factor structure in the CAPM's *assumptions.* But there is also nothing about factor structure in the CAPM's *conclusions.*" He then summarizes his position in the statement that "there can be no conflict between the CAPM and factor structure."

Later, when I was on the faculty of the Johnson School at Cornell University, I was invited to organize a conference on the impact of rate of return regulation on regulated utilities. One of my tasks was to select the keynote speaker. I turned to Jack and came to recognize his ability to move between different subjects with an ease and fluency that defied confusion and mystery. Jack with good humor challenged the audience of senior utility executives to question their preconceived notions as to both the costs/benefits of regulation and exactly the definition of what should be regulated in a highly original and thought provoking talk. Later, Jack published a similar article in the *FAJ.*[3]

For the last several years I have focused on understanding the issues related to the investments of households, their financial goals, and approaches designed to help individuals achieve the appropriate balance between investing and consumption. Unfortunately this subject engenders confusion and mystery among both clients and advisors! It turns out that both individual investors and their advisors owe a large debt to Jack, who has had an enormous influence on this group, and I and many others are fortunate to follow him. Using the platform of the *FAJ,* Jack (and his pseudonym, Walter Bagehot) made many contributions to help advisors better understand the impact of financial economics on their environment. Many of his articles have focused on active vs. passive management and their relative benefits, the appropriate use of index funds, and the cost of trading. For many years, advisors and individual investors have found it difficult to reconcile the message of thoughtful commentators such as Jack that high costs of management and the bearing of uncompensated risk can only be harmful to long-term performance in a context that stridently promotes high cost active management.

Interestingly, as the "baby boomers" move to retire and start dis-investing, the "millennial" generation has become among the wealthiest demographic sectors of the population. This demographic has grown up to be highly competent technolog-

---

[2]For a detailed description of this development see Rudd, A., & Clasing, H. K. *Modern portfolio theory: The principles of investment management* (2nd ed.). Andrew Rudd, 1988, or Grinold, R., & Kahn, R. (2000). *Active portfolio management* (2nd ed.). McGraw-Hill.

[3]Treynor, J. (2003, July–August). How to regulate a monopoly. *Financial Analysts Journal*.

ically, self-assured, and more willing to use social media to gain an understanding of what they don't know but feel they need to learn. Fittingly, many of them have discovered the multiple, lost-cost, passive strategies that have recently been delivered over automated platforms.[4] Ironically, as the investment baton is passed to a new generation of investors, they have adopted many of the strategies consistent with those advocated by Jack for a successful investment experience.

Walnut Creek, CA, USA Andrew Rudd
June 30, 2015

[4]I refer to the so-called Robo-Advisors who have gained considerable notoriety over the last few years. See the Robo-Advisor entry in Wikipedia at https://en.wikipedia.org/wiki/Robo-Advisor.

Jack Treynor: An Appreciation[1]

It is standard fare to use the hackneyed phrase "seminal article" when writing for the commemorative volume of any scholar. But when that scholar is Jack Treynor, "seminal article" is impressively ambiguous. Which one? Although this volume celebrates the 50th anniversary of Treynor's seminal article on performance measurement, and many of the chapters will focus on this one contribution, I would like to present a broader appreciation of Mr. Treynor's collection of many seminal articles.

Both practitioners and academics can claim Jack Treynor as one of their own.[2] In the course of his long and remarkable career, he has been analyst and manager, editor and professor. And as author of one of the most influential unpublished articles in the history of financial economics, Jack Treynor occupies a special place in the firmament of modern finance.

Originally a Midwesterner, Treynor was born in the small Iowa city of Council Bluffs. Although a suburb of Omaha today, Council Bluffs during Treynor's childhood was the fifth largest railroad center in the country, with eight major railroad lines passing through the city. The rail system was an important icon of twentieth century engineering and ingenuity, and Treynor's early exposure to trains—he still maintains an elaborate Lionel model train set—may well have contributed to his unique talent and taste for translating theory into practice.

Despite being born into a family of doctors, Treynor had a youthful passion for physics. None of his high school teachers had the mathematical background to help him derive the formulas in his physics textbook, so he independently invented a form of calculus to prove the equations. This aptitude for independent invention would become a hallmark of Treynor's extraordinary career. For example, he won the 1947 Westinghouse (now Intel) Science Talent Search by submitting a paper on Finite Differential Calculus—he was the winner for Iowa. But even in

[1] I thank John Guerard and Betsy Treynor for helpful comments and discussion.

[2] Biographical information taken from AFA (2007), Bernstein (1992), Mehta (2006), and Trammell (2007).

high school, it was apparent that Treynor had wide-ranging interests; his yearbook entry lists the following activities and accomplishments: "College Prep R.O.T.C. Major (highest rank), Logo (Literary Society) Spring semester President, Varsity Debate, Intersociety Debate, Football. Road Show (drama), German Club, Red Cross Council."

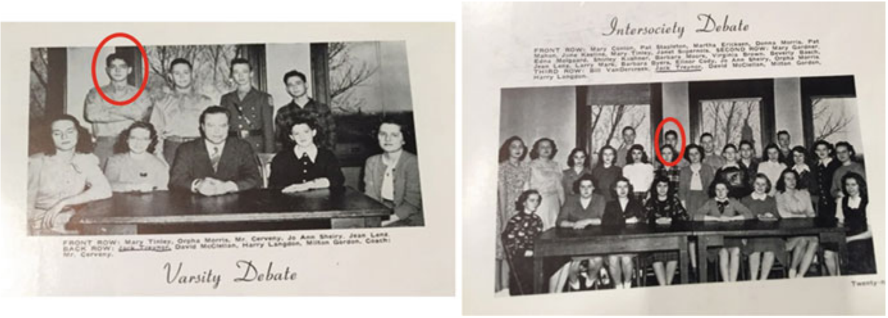


Fig. 1 Photos of Jack Treynor from his high school yearbook (courtesy of Elizabeth Treynor)

Treynor originally intended to study physics at Haverford College, but he discovered that the department only had two faculty members and became a mathematics major instead. Jack was drafted into the Army as a Private immediately upon graduating from Haverford and served from 1951 to 1953. It took some time before training and testing in Chicago was completed, after which he was sent to Monmouth, NJ, to work at the U.S. Signal Corps Laboratory for the Army. As the R&D engine of the army's Signal Corps—the division responsible for radio communications, target detection, missile guidance systems, and other sensitive military technologies—this laboratory attracted some of the army's best and brightest, as well its share of spies and defectors, including the infamous Julius Rosenberg.

Treynor served his country using his quantitative gifts and was sent to work with the man in charge of the Signal Corps, not an officer but a math Ph.D., and Treynor recalls that he "worked very hard for him." In fact, he reported to two individuals, the Ph.D. and another individual who was a military officer and periodically reminded Treynor that he was "in the Army and had orders to obey."

At one point during his service, Treynor was designated "Soldier of the Month," and although he was never told why, quite likely it was not just because of his mathematical prowess. He was an unusual person at Monmouth as he had a college education and had also trained for football. He worked digging ditches during the summers, delivered purchases from the Council Bluffs department store on his bicycle up and down the steep Bluffs in any weather, and was very fit. His Army boss had him lead miles of running and exercises and told the other soldiers they did not have to keep up with Treynor.

During this time, Treynor also attended a lecture given by Albert Einstein at Princeton's Institute for Advanced Studies—an hour's drive from Monmouth—and

Fig. 2 The U.S. Signal Corps Laboratories, Monmouth, NJ (source: http://www.armysignalocs.com/veteranssalultes/sig_corps_korean_war.html)

spoke with Einstein afterward, noting that Einstein had no ego, was more interested in asking questions, and was intellectually curious. Fortunately for the rest of us in finance, Treynor was not seduced by this remarkable encounter. Instead, he applied to Harvard Business School before leaving the army, was accepted in the Fall of 1953, and graduated with Distinction in 1955. Three professors asked Treynor to stay on at Harvard. He chose to work with the accounting expert Robert Anthony for a year, writing case studies about manufacturing companies. Anthony believed that the Boston consulting firm of Arthur D. Little would be an appropriate fit for Treynor's talents. Treynor agreed.

Treynor went on to work in the Operations Research department of ADL, where he learned of the power of the electronic computer, even becoming an early programmer, although he modestly recounts, "I wasn't good at it, because I'm not very logical."[3] During his summer vacations, he would stay with his parents at their summer home in the Colorado mountains. In 1958, on a trip to the University of Denver's library, he came across Franco Modigliani and Merton Miller's famous paper, "The Cost of Capital, Corporation Finance and the Theory of Investment." It inspired Treynor. Over the next 3 weeks of his Colorado vacation, Treynor made 44 pages of notes on a problem that had been gnawing at him since his days at Harvard Business School. Today we know Treynor's solution as the Capital Asset Pricing Model.

Many great inventions appear almost simultaneously. We think of Alexander Graham Bell and Elisha Gray inventing the telephone at the same time, or in

[3] Bernstein (1982, p. 184).

mathematics, the independent discovery of non-Euclidean geometry by János Bolyai, N.I. Lobachevsky, and Carl Friedrich Gauss. The CAPM has four nearly simultaneous inventors: William Sharpe, then at the University of Washington, whose name is most closely associated with its discovery; John Lintner at Harvard Business School; the Norwegian economist Jan Mossin; and Jack Treynor at ADL, who spent his weekends at his office composing a rough draft of his result.[4] Unknown to Treynor, a colleague in Operations Research sent a version of his draft to Merton Miller, who sent a copy to Franco Modigliani, who was moving to MIT. Modigliani read the draft and contacted Treynor, telling him over lunch that he needed to come to MIT and study economics.

Treynor took an unplanned sabbatical from ADL in 1962 to learn economics at MIT under Modigliani's supervision. Who could pass up such an opportunity? Although ADL was next-door neighbors to MIT in physical location, the two organizations were mentally thousands of miles apart. Modigliani personally chose Treynor's courses and his teachers, while Treynor picked up something of Modigliani's idiosyncratic way of looking at the world. But possibly the most important thing Modigliani did for Treynor was to break his draft in half and give it a good "social science" title. The resulting paper from the first half, "Toward a Theory of Market Value of Risky Assets," would become one of the most influential pieces of financial samizdat of the century. Twentieth-generation copies of Treynor's paper were passed around as if they contained the secrets of the Universe. Nothing quite so grandiose: they merely contained Treynor's unique formulation of the CAPM theorem, made several years before William Sharpe.

Treynor presented his paper at the MIT finance faculty seminar in the Fall of 1962 to only mild interest. Not even Modigliani was aware of its full implications. Treynor returned to ADL after his year at MIT, where his boss asked him, "Does any of this stuff that you've been doing have any commercial value at all?"[5] A few months later, Modigliani informed Treynor about Sharpe's parallel research, and suggested they exchange drafts. They did, but Treynor believed that if Sharpe was going to publish, there was little point for him to publish as well. One wonders what other treasures might lie in Treynor's file drawers?

ADL wanted commercial value from Treynor's scholarship. Treynor gave it to them in the form of two papers, "How to Rate Management of Investment Funds," which this volume commemorates, and "Can Mutual Funds Outguess the Market?" (with Kay Knight Mazuy), both of which Treynor submitted to the *Harvard Business Review* to drum up business for the firm. Not many people at ADL were interested in Treynor's papers. However, there was one new employee who was fascinated by Treynor's work: Fischer Black. Although they only overlapped at ADL for 18 months, the two men became friends, and the connection was reinforced when Black was assigned Treynor's casework after Treynor's departure for New York and Merrill Lynch in 1966. Treynor and Black collaborated on three papers, the first

[4] See French (2003) for a comparison and detailed chronology of Treynor's early work.

[5] AFA (2007, p. 12).

two on their joint model of portfolio construction, and the third a reworking of the second half of Treynor's original draft that Modigliani had cut in two at MIT.

Treynor was hired by future Treasury Secretary Don Regan to set up something new at Merrill Lynch: a computer-intensive quantitative research group. Here again, Treynor was well ahead of the curve. But Treynor found Merrill Lynch's corporate culture too limiting. "If I had stayed at Merrill, they would have made me into a narrow quant," he recounted to Peter Bernstein.[6] When the opportunity arose for him to become editor of the *Financial Analysts Journal* in 1969, he took it, even though he had never edited anything before.

As editor of the *FAJ*, Treynor found himself in a position to shape the intellectual course of an industry. Using his editorship as a bully pulpit, Treynor promoted the new quantitative ideas in financial analysis over the objections of the old guard. He also discovered he had a flair for writing. As editor, Treynor sometimes adopted the nom de plume of "Walter Bagehot," in homage to the famous nineteenth-century editor of *The Economist*. Under this name he was able to write short pieces that always cut to the heart of the matter, for example, his classic 1971 article, "The Only Game In Town," which explained the real economic role of market makers in less than four pages, anticipating by over a decade the now-standard models by Kyle (1985), Glosten and Milgrom (1985), Admati and Pfleiderer (1988), and a substantial portion of the subsequent literature in market microstructure.

Treynor left the editorship of the *FAJ* in 1981. However, even as president of his own capital management firm, Treynor still found time to pass on his ideas in academic settings. The famous illustration of the "wisdom of crowds," in which a group of students guess the correct number of jellybeans in a jar, popularized by the writer James Surowiecki, actually comes from a classroom experiment Treynor conducted at the University of Southern California. But Treynor has always rejected the role of the "narrow quant." During a period under a noncompetition agreement, Treynor wrote an award-winning play on the Kennedy assassination, and he remains fascinated by the boogie-woogie piano music of his youth. Jack Treynor's career is proof that any starting point can lead to momentous results by the inquiring mind.

Jack Treynor has been an enduring source of inspiration and intellectual leadership for me and for all my colleagues in academic finance and the financial industry, and it is a pleasure and an honor for me to be part of this wonderful and timely celebration of his contributions to performance measurement and financial economics.

MIT Sloan School of Management — Andrew W. Lo
MIT Computer Science and Artificial Intelligence Laboratory
Cambridge, MA, USA
February 10, 2015

[6] Bernstein (1992, p. 197).

References

Admati, A., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies, 1*, 3–40.

American Finance Association. (2007). Transcript of Jack Treynor interview conducted by Steve Buser. Retrieved 20 January, 2015, from http://www.afajof.org/SpringboardWebApp/userfiles/afa/file/Treynor%20transcript.doc.April13.

Bernstein, P. L. (1992). *Capital ideas: The improbable origins of modern wall street*. New York: The Free Press.

French, C. W. (2003). The treynor capital asset pricing model. *Journal of Investment Management, 2*, 60–72.

Glosten, L., & Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*, 71–100.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica, 53*, 1315–1336.

Mehta, N. (2006). FEN one on one interview: Jack Treynor. *Financial Engineering News,* 49(May/June), 5–12. Retrieved 20 January, 2015, from https://web.archive.org/web/20060904085553/http://www.fenews.com/fen49/one_on_one/one_on_one.html.

Surowiecki, J. (2003). *The wisdom of crowds*. New York: Doubleday.

Trammell, S. (2007). Making connections. *CFA Magazine, 18*, 40–44.

Treynor, J. L. (1962). Toward a theory of market value of risky assets. Unpublished manuscript.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.

Treynor, J. L. (as “Walter Bagehot”) (1971). The only game in town. *Financial Analysts Journal, 27*, 12–14+22.

Treynor, J. L. (1987). Market efficiency and the bean jar experiment. *Financial Analysts Journal, 43*, 50–53.

Treynor, J. L., & Black, F. (1972). Portfolio selection using special information, under the assumptions of the diagonal model, with mean-variance portfolio objectives, and without constraints. In G. P. Szego & K. Shell (Eds.), *Mathematical methods in investment and finance* (Vol. 4, pp. 367–384). Amsterdam: North-Holland.

Treynor, J. L., & Black, F. (1973). How to use security analysis to improve portfolio selection. *Journal of Business, 46*, 66–86.

Treynor, J. L., & Black, F. (1976). Corporate investment decisions. In S. C. Myers (Ed.), *Modern developments in financial management* (pp. 310–327). New York: Praeger.

Treynor, J. L., & Mazuy, K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131–136.

Tribute to Jack Treynor

In the early 1970s, I was working for CREF under Roger Murray, who at that time was the CIO. It was a time when investors were focusing on growth stocks or as some would characterize these as "one decision stocks": valuation had no place in investing and there was little consideration of risk. This mania captivated many investors including the various institutional investment committees.

Roger Murray was interested in bringing some discipline to investing and wanted to conduct a risk-adjusted performance study. He asked me to head up a newly created quantitative group and also retained a consultant: Fischer Black. I did not know Black but quickly found out that he was the "real deal." It became a very effective working relationship.

It turned out that Black also knew Jack Treynor very well since they had worked together on consulting projects in the past. Black initiated contact with Jack to share his expertise and he became a third party to the project. This was the beginning of a long-time relationship; I found Jack to be very knowledgeable, insightful, and creative—traits that he continued to show throughout our subsequent dealings.

In addition, I have had extensive dealings with Jack as editor of the *FAJ* journal. He is probably the best I have dealt with over my years of writing and submitting manuscripts. He could quickly recognize the worth of a manuscript and where it best fit in a journal; he would also work with me in making improvements. Two articles that Jack encouraged and helped expedite were "Homogeneous Stock Groups" and "Can Active Management Add Value," coauthored with Keith Ambachtsheer. Jack was also generous in contributing to a publication by McGraw-Hill of a textbook that I was writing: *Portfolio Management*.

One of the controversies of the day was the superiority of either the CAPM or APT. Jack contributed a superb analysis of these two theories in a chapter of my book. It was a masterful job of comparing these two and showing the equivalence.

In 1976, I was elected Chairman of the "Q" Group the Institute for Quantitative Research in Finance. At that time, my major goal was to create programs of the highest quality, ones that would focus on practical applications of economic and financial theories that had been developing over prior years.

Fortuitously, Jack had been a long-standing member of “Q”; he was also one of the pioneers in developing these theories. As editor of *FAJ*, he had broad contacts across the investment business as well as academia, so he became a source of ideas for topics as well as speakers. Within “Q”, Jack soon became an invaluable member of the research committee and has continued to contribute new ideas and suggestions for research themes and papers. As of Spring 2015, the “Q” prizes for excellence of research will be designated in honor of Jack as the Treynor Prize.

> “To live for a time close to great minds is the
> best kind of education.”
>
> John Buchan

Thanks, Jack . . .
Jim

Contents

1 **The Theory of Risk, Return, and Performance Measurement** 1
John Guerard

2 **Portfolio Theory: Origins, Markowitz and CAPM Based Selection** .. 39
Phoebus J. Dhrymes

3 **Market Timing** 49
Ravi Jagannathan and Robert A. Korajczyk

4 **Returns, Risk, Portfolio Selection, and Evaluation** 73
Phoebus J. Dhrymes and John B. Guerard

5 **Validating Return-Generating Models** 111
Marshall E. Blume, Mustafa N. Gültekin, and N. Bülent Gültekin

6 **Invisible Costs and Profitability** 135
Xiaoxia Lou and Ronnie Sadka

7 **Mean-ETL Portfolio Construction in US Equity Market** 155
Barret Pengyuan Shao

8 **Portfolio Performance Assessment: Statistical Issues and Methods for Improvement** 169
Bernell K. Stone

9 **The Duality of Value and Mean Reversion** 229
Noah Beck, Shingo Goto, Jason Hsu, and Vitali Kalesnik

10 **Performance of Earnings Yield and Momentum Factors in US and International Equity Markets** 239
Jose Menchero and Zoltán Nagy

11 **Alpha Construction in a Consistent Investment Process** 257
Sebastián Ceria, Kartik Sivaramakrishnan, and Robert A. Stubbs

12 Empirical Analysis of Market Connectedness as a Risk Factor for Explaining Expected Stock Returns 275
Shijie Deng, Min Sim, and Xiaoming Huo

13 The Behaviour of Sentiment-Induced Share Returns: Measurement When Fundamentals Are Observable 291
Richard A. Brealey, Ian A. Cooper, and Evi Kaplanis

14 Constructing Mean Variance Efficient Frontiers Using Foreign Large Blend Mutual Funds 315
Ganlin Xu, Harry Markowitz, Minyee Wang, and John B. Guerard

15 Fundamental Versus Traditional Indexation for International Mutual Funds: Evaluating DFA, WisdomTree, and RAFI PowerShares 331
Heehyun Lim and Edward Tower

16 Forecasting Implied Volatilities for Options on Index Futures: Time-Series and Cross-Sectional Analysis versus Constant Elasticity of Variance (CEV) Model 355
Tzu Tai and Cheng Few Lee

17 The Swiss Black Swan Unpegging Bad Scenario: The Losers and the Winners 389
Sebastien Lleo and William T. Ziemba

18 Leveling the Playing Field 421
Jonathan B. Berk and Jules H. van Binsbergen

19 Against the 'Wisdom of Crowds': The Investment Performance of Contrarian Funds 431
Kelsey D. Wei, Russ Wermers, and Tong Yao

Author Bios

Noah Beck is a Senior Researcher at Research Affiliates. Noah conducts quantitative equity research relating to dynamic factor investing and smart beta strategies. He supports existing portfolios and advances research for product development. Noah received a B.S. in physics from Harvey Mudd College and a Master of Financial Engineering from the Anderson School of Management at UCLA.

Jonathan Berk is the A.P. Giannini Professor of Finance at the Graduate School of Business, Stanford University. His research interests cover a broad range of topics in finance including delegated money management; the pricing of financial assets; valuing a firm's growth potential, the capital structure decision, and the interaction between labor markets and financial markets. His work is internationally recognized and has won numerous awards, including the TIAA-CREF Paul A. Samuelson Award, the Smith Breeden Prize, Best Paper of the Year in the *Review of Financial Studies*, and the FAME Research Prize. His article, "A Critique of Size Related Anomalies," was selected as one of the two best papers ever published in the *Review of Financial Studies*, and was also honored as one of the 100 seminal papers published by Oxford University Press. In recognition of his influence on the practice of finance he has received the Graham and Dodd Award of Excellence, the Roger F. Murray Prize, and the Bernstein-Fabozzi/Jacobs Levy Award. He served as an Associate Editor of the *Journal of Finance* from 2000 to 2008, is currently an associate editor of the *Journal of Portfolio Management*, and is a Research Associate at the National Bureau of Economic Research. He served a term on the board of directors of the Financial Management Association and the American Finance Association.

Jonathan Berk received his Ph.D. in Finance from Yale University. Before coming to Stanford he was the Sylvan Coleman Professor of Finance at Haas School of Business at the University of California, Berkeley.

Richard Brealey is Emeritus professor of Finance at the London Business School. He is a former president of the European Finance Association a former director of the American Finance Association. He is a fellow of the British Academy. He has served as a special advisor to the Governor of the Bank of England and as a director of a number of financial institutions. His publications include Principles of Corporate Finance (with Stewart Myers and Franklin Allen, 11th ed. 2014).

Sebastian Ceria, Ph.D., is Chief Executive Officer of Axioma. Before founding Axioma, Ceria was an Associate Professor of Decision, Risk, and Operations at Columbia Business School from 1993 to 1998. Ceria has worked extensively in the area of optimization and its application to portfolio management. He is the author of many articles in publications including Management Science, Mathematical Programming, Optima, and Operations Research. Most recently, Ceria's work has focused on the area of robust optimization in portfolio management. He has coauthored numerous papers on the topic, including, "Incorporating Estimation Errors into Portfolio Selection: Robust Portfolio Construction," published in *The Journal of Asset Management*, "To Optimize or Not to Optimize: Is That the Question?" published in the *Oxford Handbook of Quantitative Management*, and "Factor Alignment Problems and Quantitative Portfolio Management," published in the *Journal of Portfolio Management*. He is a recipient of the Career Award for Operations Research from the National Science Foundation. Ceria completed his Ph.D. in Operations Research at Carnegie Mellon University's Tepper School of Business, and his undergraduate degree in Applied Math at the University of Buenos Aires, Argentina.

Ian Cooper is Professor of Finance at London Business School. He has held positions at the University of Chicago and the Australian Graduate School of Management and has served as a consultant for a number of major institutions. His research covers international capital markets, portfolio investment, and behavioral finance.

Shijie Deng, Ph.D., is an Associate Professor of industrial and systems engineering at Georgia Institute of Technology, Atlanta. His research interests include financial asset pricing and real options valuation, financial engineering applications in energy commodity markets, transmission pricing in electric power systems, stochastic modeling, and simulation. He has served as a consultant to several private and public organizations on issues of risk management and asset valuation in the restructured electricity industry.

Phoebus J. Dhrymes was the Edwin W. Rickert Professor of Economics Emeritus at Columbia University. Professor Dhrymes passed away while his chapter was in galley proofs. He received his B.A. degree (with Highest Honors) from the University of Texas (Austin) in 1957, and his Ph.D. in economics from the Massachusetts Institute of Technology in 1961.

He taught previously at MIT, Harvard University, the Wharton School University of Pennsylvania and was a visiting scholar or professor at Stanford University, the University of California at Los Angeles, Osaka University (Japan), Monash University (Australia), and the Institute for Advanced Studies in Vienna (Austria).

He was an editor of the *International Economic Review*, a Founding Editor of the *Journal of Econometrics*, and was a Fellow of the Econometric Society and the American Statistical Association. He was the author of several monographs, including *Econometrics: Statistical Foundations and Applications* (1970), *Distributed Lags: Problems of Formulation and Estimation* (1971), *Introductory Econometrics* (1977), *Topics in Advanced Econometrics*, vols. I and II, (1989, 1994), *Time Series Unit Roots and Cointegration* (1998), and *Mathematics for Econometrics*, fourth edition (2013). He published numerous articles in, among others, *Econometrica*, *Journal of Econometrics*, *Journal of the American Statistical Association*, *The Australian Journal of Statistics*, *The American Economic Review*, and the *Journal of Finance*.

James L. Farrell, Jr. is the Chairman of the Institute for Quantitative Research in Finance (the "Q" Group), one of the premier investment forums in the world. He is on the Board of Trustees of Capital Guardian Funds. He was Managing Director of Research and Development at Ned Davis Research. Before joining the firm, Dr. Farrell was Managing Director of Morse Williams, a New York City-based investment advisory firm. He also founded Farrell-Wako Global Investment Management, a joint venture with Wako Securities, a mid-sized brokerage firm, which subsequently merged with Sumitomo Life (SL). Dr. Farrell also worked as a portfolio manager at TIAA-CREF, Citibank, and MPT Associates. He has been published in investment journals, such as the *Journal of Portfolio Management* and *Financial Analysts Journal* and is the author of several textbooks on portfolio management including *Portfolio Management: Theory and Application*. His writings and lectures cover asset allocation, security valuation, portfolio analysis, international investing, and the merits of introducing greater discipline and structure in the investment process. He holds a Ph.D. in finance and economics from New York University, an M.B.A. from the Wharton School of Business at the University of Pennsylvania, and a B.S. degree from the University of Notre Dame. He also holds the Chartered Financial Analyst (CFA) designation.

Shingo Goto, Ph.D., is a clinical associate professor of finance at the Moore School of Business, University of South Carolina. His research has been published in the *Review of Financial Studies*, *the Journal of Financial and Quantitative Analysis*, and *the Journal of Banking and Finance*, among others. He was a principal at the Active Equity group of Barclays Global Investors, and has been a consultant to Research Affiliates and Nomura Asset Management. He holds a Ph.D. and M.B.A. from the Anderson Graduate School of Management at UCLA and a bachelor's degree from the University of Tokyo.

John B. Guerard, Jr., Ph.D., is Director of Quantitative Research at McKinley Capital Management, in Anchorage, Alaska. He earned his AB in Economics from Duke University, M.A. in Economics from the University of Virginia, MSIM from the Georgia Institute of Technology, and Ph.D. in Finance from the University of Texas, Austin. John taught at the McIntire School of Commerce, the University of Virginia, Lehigh University, and Rutgers University. John taught as an adjunct faculty member at the International University of Monaco and the University of Pennsylvania. He worked with the DAIS Group at Drexel, Burnham, Lambert, Daiwa Securities Trust Company, Vantage Global Advisors, and served on the

Virtual Research team at GlobeFlex Capital. John co-managed a Japanese equity portfolio with Harry Markowitz at Daiwa Securities Trust Company. While serving as Director of Quantitative Research at Vantage Global Advisors (formerly MPT Associates), Mr. Guerard was awarded the first Moskowitz Prize for research in socially responsible investing. Mr. Guerard has published several monographs, including *Corporate Financial Policy and R&D Management* (Wiley, 2006, second edition), *Quantitative Corporate Finance* (Springer, 2007, with Eli Schwartz), *The Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques* (Springer, 2010), and *Introduction to Financial Forecasting in Investment Analysis* (Springer, 2013). John serves an Associate Editor of the *Journal of Investing* and *The International Journal of Forecasting*. Mr. Guerard has published research in *The International Journal of Forecasting*, *Management Science*, the *Journal of Forecasting*, *Journal of Investing*, *Research in Finance*, the *IBM Journal of Research and Development*, *Research Policy*, and the *Journal of the Operational Research Society*.

Nihat Bülent Gültekin has been on the faculty of the Finance Department of the Wharton School of the University of Pennsylvania since 1981. He was the Academic Director for the Wharton MENA Center in Abu Dhabi during 2011–2013.

He taught at the Amos Tuck School at Dartmouth College and University of Chicago before he joined the Wharton School. He was a visiting professor of finance at Koc University, INSEAD, Bosporus University, and University Paris at Dauphine. He received his B.Sc. in mechanical engineering from Robert College, M.B.A. from Bogazici University, and Ph.D. from the Wharton School.

He held the following official positions in Turkey: Governor of the Central Bank of the Republic of Turkey; Chief Advisor to Prime Minister Turgut Özal; Undersecretary of Mass Housing and Public Participation Administration, a state agency in charge of housing and urban development, large-scale infrastructure projects, and the privatization program of the Turkish Government; Director General for Research and Planning at the Central Bank of the Republic of Turkey.

He also served as an advisor for Turgut Ozal when he became president.

He served as the Chief Advisor to the Minister of Privatization in Poland during 1989–1991. He was a policy advisor to the governments of Indonesia, Egypt, Uzbekistan, Kazakhstan, Ukraine, Belarus, Russian Federation, and to the presidents Nursultan Nazarbayev and Islam Karimov.

He served as a senior advisor to the many agencies in the public sector such as USAID, European Bank of Restructure and Development, International Finance Corporation, the World Bank, Istanbul Stock Exchange, OECD, UNDP, Capital Markets Board of Turkey, and Institute of Banking of the Saudi Monetary Agency. He was a consultant for many international firms including Charles River Associates, Goldman Sachs, Morgan Stanley & Co., Rothschild & Cie, IBM, Merck and Co., Citicorp, Chemical Bank, Barclays Bank, Merrill Lynch, Namura Securities, Arab National Bank in Riyadh, and the Dubai World. He was the Chairman of the Board of Sumerbank, a state owned conglomerate with 35,000 employees and a director on the boards of Emlak Bank, Westergaard Fund, and Bell Atlantic Mutual Funds.

He is currently on the boards of Limak Yatirim in Turkey, QFS Funds in the USA, and Oasis Capital in Geneva.

He is a founding trustee of the Koc University, Foundation of Ownership Changes in Warsaw, a former trustee of the American University in Bulgaria. International Business School in Moscow, and International Institute of Business in Kiev.

Prof. Gultekin published extensively in scientific journals.

Mustafa N. Gültekin Investments, portfolio theory, asset pricing models, financial modeling, valuation, and risk management are the focus of Prof. Mustafa N. Gültekin's work. He teaches applied investment management, investments, financial modeling, valuation and corporate restructuring, and financial markets. Other areas of expertise include Monte Carlo simulations, international finance, mortgage backed securities (MBS), and asset-liability management.

Dr. Gültekin has served as a consultant to major corporations in the United States and abroad. He is a limited partner at the Blackethouse Group LLC, a restructuring consulting group, senior advisor to Morning Meeting Inc., a financial modeling and consulting group, and served as a long-time consultant to the Community First Investment Risk Evaluation team (CFIRE) of Community First Financial Group. At CFIRE he led a research team building MBS valuations and stochastic interest rate models to manage the asset-liability portfolios of the Community First Financial Groups' banks. He worked with the Economic Development Group of the Kenan Institute of Private Enterprise in building financial models for evaluation and feasibility of alternative energy sources.

Dr. Gültekin is the former president of the European Financial Management Association and the former dean of the College of Administrative Sciences and Economics at Koç University in Istanbul. He also served as associate director of the Management Decision Laboratory at New York University and as a research scientist at Boğazici University in Turkey.

He received his Ph.D. in Finance from New York University, his M.A. in Operations Management from Boğazici University, Istanbul, Turkey, and a B.S. in physics from Middle East Technical University, Ankara, Turkey. Prof. Gültekin has been on the Finance Faculty of UNC-Chapel Hill's Kenan-Flagler Business School since 1985.

Dr. Gültekin is an avid skier and a competitive sailor. He lives in Chapel Hill, NC, with his wife Mabel Miguel, also a professor at Kenan-Flagler.

Jason Hsu is at the forefront of the Smart Beta revolution, beginning with his pioneering work with Rob Arnott on the RAFI® Fundamental Index® approach. He has published numerous articles and is coauthor of *The Fundamental Index: A Better Way to Invest*. Jason served as the firm's CIO from 2005 to 2014, and in his current role as Vice Chairman he is in charge of leading the firm's long-term strategic initiatives and shaping the firm's vision and culture.

In addition, Jason is an adjunct professor in finance at the Anderson School of Management at UCLA and a visiting professor in international finance at the Taiwan National University of Political Science.

Jason began his career working in derivatives research and in the trading area of Taiwan's Far Eastern Securities. He has also been a consultant to Asian investment banks, securities dealers, and insurance companies on issues such as risk management, strategic asset allocation, and convergence trading.

Jason graduated summa cum laude with a B.S. in physics from the California Institute of Technology, was awarded an M.S. in finance from Stanford University, and earned his Ph.D. in finance from UCLA.

Jason Hsu, Ph.D., is Founder and Chief Investment Officer at Rayliant Global Advisors, an Asian specialist asset manager, and Co-Founder and Vice Chairman at Research Affiliates, the leading provider of Smart Beta indexes. Jason is best known for his research on Smart Beta, Factor Investing, Manager Selection, and Asset Allocation. He has published more than 40 articles in academic and practitioner journals and contributed to six handbooks on quantitative investing. Jason's research have won 2 Bernstein Fabozzi/Jacob Levy Outstanding Article awards, 2 William Sharpe Best Index Research Award, 1 Graham and Dodd Scroll Award, and 1 FAJ Readers' Choice Award. Jason holds a Ph.D. from UCLA Anderson School of Management, where he is also currently an adjunct professor in finance. Jason has also held visiting professorship at UC Irvine Merage Business School, Taiwan National Chengchi University, and Tokyo Metropolitan University.

Xiaoming Huo, Ph.D., received his B.S. in mathematics from the University of Science and Technology, China, in 1993 and his M.S. in electrical engineering and his Ph.D. in statistics from Stanford University, Stanford, CA, in 1997 and 1999, respectively. He is a Professor with the School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta. His research interests include dimension reduction and nonlinear methods. Dr. Huo received first prize in the Thirtieth International Mathematical Olympiad (IMO), which was held in Braunschweig, Germany.

Ravi Jagannathan is the Chicago Mercantile Exchange/John F. Sandner Professor of Finance at Northwestern University's Kellogg School of Management and Co-Director of the Financial Institutions and Markets Research Center at the Kellogg School (1997—present). He has previously held positions as Piper Jaffray Professor of Finance (1993–1997) and Associate Professor of Finance (1989–1993) at the University of Minnesota's Carlson School of Management, Assistant Professor of finance at Northwestern University's Kellogg School (1983–1989), and as a Distinguished Visiting Professor at the Hong Kong University of Science and Technology (1994–1995). Ravi received a Ph.D. in Financial Economics (1983) and an M.S. in Financial Economics (1981) from Carnegie Mellon University, an M.B.A. from the Indian Institute of Management at Ahmedabad (1972), and a B.E. in Mechanical Engineering from the University of Madras (1970). Ravi has served on the editorial boards of leading academic journals and is a former executive editor of the *Review of Financial Studies*. He has served as a member of the Board of Directors of the American Finance Association and the Western Finance Association and is a past President of the Western Finance Association, the

Society of Financial Studies, and the Financial Intermediation Research Society. He is a research associate of the National Bureau of Economics Research, a fellow of the Society for Financial Econometrics, and a member of the Board of Directors of the Financial Management Association. He is a special term professor at the Indian School of Business and the Shanghai Advanced Institute of Finance.

Ravi's research interests are in the areas of asset pricing, capital markets, and financial institutions. His articles have appeared in leading academic journals, including the *Journal of Political Economy, Journal of Financial Economics, Journal of Finance*, and *Review of Financial Studies*. He is noted for his contributions to the development of the Hansen-Jagannathan bound, and the Hansen-Jagannathan distance that summarizes what is missing in an asset pricing model, the TGARCH volatility model, a contingent claims framework for evaluating the performance of actively managed portfolios, and the role of portfolio weight constraints in estimating vast covariance matrices with precision.

Vitali Kalesnik, Ph.D., is Director and Head of Equity Research at Research Affiliates. Vitali is responsible for quantitative research using advanced econometric tools in asset pricing and active asset allocation. This research is used to enhance Research Affiliates products—in particular, RAFI™ Fundamental Index™ strategies and global tactical asset allocation products. Vitali earned his Ph.D. in economics from the University of California, Los Angeles, where he was a winner of the UCLA Graduate Division Fellowship for 2001–2005.

Evi Kaplanis is an Honorary Senior Research Fellow at the London Business School. She was an Associate Professor of Finance at LBS until 2000. She has published widely in the area of International Finance in academic journals including the *Review of Financial Studies* and the *Journal of International Money and Finance*. She taught across all the programs of the School and was the first director of the Full-time Masters Program in Finance at LBS. She has also served as a consultant for a number of financial institutions.

Robert A. Korajczyk a member of the Kellogg School faculty since 1982. At Kellogg, Korajczyk has previously served as Senior Associate Dean: Curriculum and Teaching, Chair of the Department of Finance, Director of the Zell Center for Risk Research.

Professor Korajczyk's research interests are in the areas of investments and empirical asset pricing. He is a recipient of the 2009 Crowell Prize for best paper in the field of quantitative asset management, awarded by PanAgora Asset Management; the Alumni Choice Faculty Award 2000; the Core Teaching Award 1998 and 2000; the Sidney J. Levy Teaching Award 1996; the New York Stock Exchange Award for Best Paper on Equity Trading, presented at the 1993 Western Finance Association annual meetings; and the Review of Financial Studies Best Paper Award, 1991.

Professor Korajczyk is a past editor of the *Review of Financial Studies* and a past associate editor of the *Review of Financial Studies*, *Journal of Business & Economic Statistics*, *Journal of Empirical Finance*, and the *Journal of Financial and Quantitative Analysis*.

He has held visiting faculty appointments at the University of Chicago, the University of Melbourne, the University of Vienna, and the Hong Kong University of Science and Technology. Professor Korajczyk serves on the product development committees of Select Innovation Investments, LLC and DSC Quantitative Group, LLC.

Professor Korajczyk received his B.A., M.B.A., and Ph.D. degrees from the University of Chicago.

Cheng-Few Lee is a Distinguished Professor of Finance at Rutgers Business School, Rutgers University, and was chairperson of the Department of Finance from 1988 to 1995. He has also served on the faculty of the University of Illinois (IBE Professor of Finance) and the University of Georgia. He has maintained academic and consulting ties in Taiwan, Hong Kong, China, and the United States for the past three decades. He has been a consultant to many prominent groups including the American Insurance Group, the World Bank, the United Nations, The Marmon Group Inc., Wintek Corporation, and Polaris Financial Group.

Professor Lee founded the *Review of Quantitative Finance and Accounting* (RQFA) in 1990 and the *Review of Pacific Basin Financial Markets and Policies* (RPBFMP) in 1998, and serves as managing editor for both journals. He was also a co-editor of the *Financial Review* (1985–1991) and the *Quarterly Review of Economics and Finance* (1987–1989). In the past 39 years, Dr. Lee has written numerous textbooks ranging in subject matters from financial management to corporate finance, security analysis and portfolio management to financial analysis, planning and forecasting, and business statistics. In addition, he edited two popular books, *Encyclopedia of Finance* (with Alice C. Lee) and *Handbook of Quantitative Finance and Risk Management* (with Alice C. Lee and John Lee). Dr. Lee has also published more than 200 articles in more than 20 different journals in finance, accounting, economics, statistics, and management. Professor Lee was ranked the most published finance professor worldwide during the period 1953–2008.

Professor Lee was the intellectual force behind the creation of the new Masters of Quantitative Finance program at Rutgers University. This program began in 2001 and has been ranked as one of the top ten quantitative finance programs in the United States. These top ten programs are located at Carnegie Mellon University, Columbia University, Cornell University, New York University, Princeton University, Rutgers University, Stanford University, University of California at Berkley, University of Chicago, and University of Michigan.

Heehyun Lim is a business consultant in the Republic of Korea.

Sébastien Lleo is an Associate Professor in the Finance Department at NEOMA Business School in France and a tutor on the Certificate in Quantitative Finance at FitchLearning in the UK. He currently serves as Director of NEOMA's doctoral programs, where he heads the Ph.D. Program and the DBA Program established in partnership with Shanghai Jiaotong University (China). He is also the lead researcher on the RISK PEFORM project, jointly funded by Région Champagne Ardennes and the European Union, a co-Director of the GARP University Chapter, and serves on the Steering Group of the CQF Institute. Sébastien was previously Research Associate at Imperial College London in the UK, worked for 7 years in the investment industry in Canada, and consulted on risk management and asset allocation projects in Canada and the UK. He also held a visiting position at the Frankfurt School of Finance and Management in Germany. He holds a Ph.D. in Mathematics from Imperial College London (UK), an M.B.A. from the University of Ottawa (Canada), and an MSc in Management from Reims Management School (France). He is also a CFA Charterholder, a Certified Financial Risk Manager, a Professional Risk Manager, and a CQF alumnus.

Andrew Lo is the Charles E. and Susan T. Harris Professor, MIT Sloan School of Management; director, MIT Laboratory for Financial Engineering; and principal investigator, MIT Computer Science and Artificial Intelligence Laboratory.

Professor Lo has published numerous articles in finance and economics journals. He is a coauthor of *The Econometrics of Financial Markets, A Non-Random Walk Down Wall Street*, *The Heretics of Finance*, and *The Evolution of Technical Analysis*, and is the author of *Hedge Funds: An Analytic Perspective*. Lo is currently an associate editor of the *Financial Analysts Journal*, the *Journal of Portfolio Management*, the *Journal of Computational Finance*, and *Quantitative Finance,* and is a co-editor of *Annual Review of Financial Economics.*

Professor Lo holds a B.A. in economics from Yale University as well as an A.M. and a Ph.D. in economics from Harvard University.

Xiaoxia Lou is an associate professor of Finance at the University of Delaware. She received her B.A. in economics from Peking University (China), her M.S. in statistics from Iowa State University, and her Ph.D. in finance from the University of Washington.

Xiaoxia's research interests are in the areas of investments, empirical asset pricing, market microstructure, and corporate finance. She has published in top finance journals such as the *Journal of Finance*, the *Journal of Financial Economics*, the *Journal of Financial and Quantitative Analysis*, and the *Financial Analyst Journal*. Her research has been presented at numerous conferences and institutions including the American Finance Association Annual Meetings (AFA), the European Finance Association Annual Meetings (EFA), the National Bureau of Economic Research (NBER), Goldman Sachs Asset Management, the Center for Research in Security Prices (CRSP) Forum, the Q group, and State Street Global Advisors. She is a recipient of the 2008 Q Group Research Grant and the Best Paper Award presented at 2008 CRSP forum at the University of Chicago.

Harry Markowitz is the Chief Architect at Guidedchoice.com in San Diego, CA, and a quantitative research advisor at McKinley Capital Management LLC in Anchorage, AK. hmarkowitz@mckinleycapital.com. 8910 University Center Lane, Suite 400, San Diego, CA 92122. He is a Noble Laureate and the father of Modern Portfolio Theory, named "Man of the Century" by *Pensions and Investments* magazine. He is a recipient of the prestigious John von Neumann Theory Prize for his work in portfolio theory, sparse matrix techniques, and the SIMSCRIP programming language. Dr. Markowitz earned his Ph.D. from the University of Chicago. He is a quantitative research scientific advisor at McKinley Capital Management LLC in Anchorage, AK.

Jose Menchero serves as Head of Portfolio Analytics Research at Bloomberg. Jose and his team are responsible for developing the full suite of factor risk models spanning multiple asset classes, as well as portfolio risk and return attribution, portfolio construction, and portfolio optimization.

Prior to joining Bloomberg, Jose was the founder and CEO of Menchero Portfolio Analytics Consulting. Before founding his consulting firm, Jose worked for 8 years at MSCI, where he was Managing Director responsible for building the next-generation suite of Barra equity risk models, and for portfolio construction research. Jose also served for 7 years as Director of Research at Thomson Financial, where he developed several risk and return attribution methodologies, as well as equity factor risk models.

Jose has over 30 finance publications in leading practitioner journals. Before entering finance, Jose was Professor of Physics at the University of Rio de Janeiro in Brazil. Jose holds a Ph.D. in theoretical physics from the University of California at Berkeley, and a BS degree in aerospace engineering from the University of Colorado at Boulder. Jose is a CFA Charterholder and was inducted into the Performance and Risk Hall of Fame class of 2014, together with Nobel Laureates Eugene Fama and Harry Markowitz.

Zoltán Nagy is a Vice President in the Equity Applied Research group at MSCI, based in Budapest, Hungary. In his current role, he focuses on investment problems arising in factor investing, with a special focus on the use of fundamental equity risk models in portfolio construction and analysis. He is also keen on analyzing the quantitative implications of integrating ESG considerations into the portfolio management process. He joined MSCI in 2008 as an Index Researcher working on the development of new index products. Prior to working in finance, Mr. Nagy was a researcher in mathematical physics at the Mathematics Department of the University of the Algarve in Faro, Portugal, with special interest in the study of quantum integrable systems and the underlying algebraic structures. Mr. Nagy holds a Ph.D. in theoretical physics from the University of Cergy-Pontoise, France, and an engineering degree from the Ecole Polytechnique, France. Mr. Nagy is also a CFA charterholder.

Andrew Rudd is an expert in quantitative investment analysis, particularly in the areas of asset allocation, asset-liability management, modern portfolio theory, and risk management.

A serial entrepreneur, Andrew has founded several finance and technology firms. He was a co-founder and former chairman and CEO of Barra, Inc., where he served as CEO from 1984 to 1999. He founded Advisor Software Inc. in 1995, with the goal of delivering world class analytics to the retail financial services market, where he serves as its Chairman and CEO.

Prior to his career as an entrepreneur, Andrew was a Professor of Finance and Operations Research at Cornell University in Ithaca, New York. In addition, he has written numerous journal articles and research papers on a wide range of domestic and international investment practices and theories. He has focused much of his recent research on Goal-Based Investing, an extension of Asset-Liability Management to individual, taxable portfolios. Some of this work was published in the Fall 2013 issue of the *Journal of Wealth Management* in an article coauthored with Larry Siegel, entitled "Using an Economic Balance Sheet for Financial Planning."

He is the coauthor of two industry-leading books on institutional investing: *Modern Portfolio Theory: The Principles of Investment Management* (1982) and *Option Pricing* (1983). He is also co-editor with Professor Stephen Satchell from Cambridge University of a collection of articles entitled "Quantitative Approaches to High Net Worth Investment," published by Risk Books in 2014. Andrew is also widely credited for some of the foundational quantitative investment analysis of the movie industry, summarized in "Investing in Movies" coauthored with Mark Ferrari and published in the *Journal of Asset Management*, May 2008.

Andrew received his Bachelor of Science degree with honors in Mathematics and Physics from Sussex University in England, and earned a M.Sc. in Operations Research, an M.B.A in Finance and International Business, and a Ph.D. in Operations Research from the University of California, Berkeley. He also serves as a Trustee of the University of California, Berkeley Foundation, and the Richard C. Blum Center for Developing Economies and as a member of the Investment Committee of the University of Massachusetts Foundation.

Ronnie Sadka is chairperson of the finance department at the Carroll School of Management, Boston College. His research focuses on liquidity in financial markets and big-data applications. Sadka's work has appeared in various outlets including the *Journal of Finance*, the *Journal of Financial Economics*, and *Financial Analysts Journal*, and has been covered by *The New York Times*, *The Wall Street Journal*, and *CNBC*. Prior academic experience includes teaching at the University of Chicago (Booth), New York University (Stern), Northwestern University (Kellogg), and the University of Washington (Foster). Industry experience includes Goldman Sachs Asset Management and Lehman Brothers (quantitative strategies). Sadka recently served on the economic advisory board of NASDAQ OMX. Professor Sadka earned a B.Sc. (Magna Cum Laude) in industrial engineering and an M.Sc. (Summa Cum Laude) in operations research, both from Tel-Aviv University. He received a Ph.D. in finance from Northwestern University (Kellogg).

Barret Pengyuan Shao, Ph.D., Crabel Capital Management LLC, 414 East Market Street, Suite I, Charlottesville, VA, 22902. Mr. Shao is Associate Portfolio Manager at Crabel Capital Management. He earned his B.S. in automatic control from Xiamen University, B.S. in mathematical economics from Wangyanan Institute for Studies in Economics (WISE) in China, and Ph.D. in applied mathematics and statistics (quantitative finance) from Stony Brook University.

Min Sim, Ph.D., is Vice President of Passive Investment Strategy team at Samsung Asset Management. His main responsibility includes developing investment solutions and offering passive investment instruments such as index funds and ETFs to investors accordingly. Prior to joining Samsung Asset Management, Dr. Sim earned his master's degree in Financial Mathematics from the University of Chicago and his doctoral degree in Industrial Engineering from Georgia Institute of Technology. His research interests include stochastic models and graphical models applied to financial asset pricing, risk management, and high-frequency financial data analysis.

Kartik Sivaramakrishnan is a Senior Researcher at Axioma where he maintains Axioma's flagship portfolio optimizer that offers the most flexible construction tool in the market. Kartik develops second-order mixed-integer portfolio models, and interior point algorithms and associated software for solving these models in a high performance computing environment. He also researches the applications of these models in improving the investment process in practice. His current research interests are in extending the equity MVO framework to (a) multi-period models in portfolio transition management, alpha decay, and tax optimization; and (b) scenario based portfolio construction methodologies for portfolios containing "multi-asset" instruments including equities, fixed-income, commodities, credit, foreign exchange, derivatives, and alternatives including private equity and hedge funds.

Prior to joining Axioma, Kartik was a tenure-track Assistant Professor in the Department of Mathematics and affiliated with the Operations Research program at North Carolina State University in Raleigh. Kartik has a Ph.D. in Applied Mathematics from Rensselaer Polytechnic Institute and a Master in Electrical Engineering degree from the Indian Institute of Science in Bangalore. His research has been published in optimization, OR, and financial practitioner journals.

Bernell K. Stone was the Harold F. Silver Professor of Finance, at Brigham Young University. He is now Professor, emeritus. He has degrees from Duke University (B.S., Physics), University of Wisconsin (M.S., Physics), and M.I.T. (Applied Mathematics and Management). He taught at Cornell, the Georgia Institute of Technology, and was founding editor of the *Journal of Cash Management* before joining the faculty at BYU. He was editor, co-editor, and associate editor of eight journals, including the *Journal of Financial and Quantitative Analysis (JFQA)* and *Financial Management*. He is the author of five books and more than 100 articles.

Robert A. Stubbs is Vice President, Strategic Innovation at Axioma, Inc. Stubbs joined Axioma in 2000 and is currently the head of the recently launched Innovation Lab at Axioma. Prior to his most recent position, Stubbs was the head of research at Axioma for more than ten years. In this role, Stubbs developed a robust portfolio optimization technology and made numerous contributions to classical portfolio optimization techniques such as developing a highly effective, specialized algorithm for solving portfolio rebalancing problems with a limit on the number of assets in the portfolio. Stubbs led the research on Axioma's equity risk models. More recently, Stubbs has focused on the interaction of equity risk models, performance attribution, and portfolio construction in the overall investment process. In these research efforts, Stubbs led the research and development of the Alpha Alignment Factor to adjustment for misalignment of factors in alpha and risk models, Risk Model Machine to generate customized risk models, and Consistent Portfolio Management to better realize the performance of alpha models.

Stubbs holds a Ph.D. in Industrial Engineering and Management Science from Northwestern University and a Bachelor of Science in Applied Mathematics from Auburn University. In his doctoral dissertation, Stubbs designed new algorithms for solving large-scale nonlinear integer programming problems. He has numerous publications in optimization and financial practitioner journals.

Edward Tower is a professor of economics at Duke University and a long-term visiting professor at Chulalongkorn University in Bangkok, Thailand. Professor earned his undergraduate degree in Physics and Ph.D. in Economics from Harvard University.

Ming Yee Wang is Chief Investment Officer at GuidedChoice.com. Ming oversees development of investment strategies and software, analyzes and manages portfolio performance, and serves on the Investment Committee. Prior to joining GuidedChoice he served in leadership and analytical roles at Investment Research Company (IRC) and TIAA-CREF Investment Management, Inc. Ming received his M.A. in Physics from Princeton University and his M.A. in Mathematics from Courant Institute. Ming has published in *Financial Analyst Journal* an article on the applications of Quadratic Optimization to portfolios with dual benchmarks.

Kelsey Wei is an associate profess of finance at the Jindal School of Management at the University of Texas at Dallas. Her current research focuses on issues related to investment management including institutional investors, mutual funds and hedge funds, financial analysts, insider trading, and credit ratings. She has also conducted research on important corporate finance issues such as the board of directors, corporate governance, and M&As. Her research has been published in leading finance and accounting journals, such as *Journal of Finance*, *Journal of Financial Economics*, *Journal of Accounting Research*, and *Management Science*. She was an economist at the Board of Governors of the Federal Reserve System. Prior to joining the University of Texas at Dallas she was an assistant professor of finance at the State University of New York at Binghamton. Kelsey Wei received a B.S. in engineering economics from Tianjin University, a M.S. in economics from Texas A&M University, and a Ph.D. in finance from the University of Texas at Austin.

Ganlin Xu is the Chief Technology Officer of Guidedchoice.com, gxu@ mckinleycapital.com, where he leads a team of finance, economics, and mathematics experts to develop Guidedchoice's financial methodology, software, and systems. His address is Guided Choice, 8910 University Lane Center #750, San Diego, 92122. Dr. Xu earned his Ph.D. from Carnegie Mellon University. His research on portfolio theory has been published in various journals. He is a quantitative research scientific advisor at McKinley Capital Management LLC in Anchorage, AK.

William T. Ziemba is the Alumni Professor (Emeritus) of Financial Modeling and Stochastic Optimization in the Sauder School of Business, University of British Columbia, where he taught from 1968 to 2006. His Ph.D. is from the University of California, Berkeley. He currently teaches part time and makes short research visits at various universities. Recently he is the Distinguished Visiting Research Associate, Systemic Risk Centre, London School of Economics.

He has been a visiting professor at Cambridge, Oxford, London School of Economics, University of Reading and Warwick in the UK, at Stanford, UCLA, Berkeley, MIT, University of Washington and Chicago in the USA, Universities of Bergamo, Venice and Luiss in Italy, the Universities of Zurich, Cyprus, Tsukuba (Japan), KAIST (Korea), and the National University and the National Technological University of Singapore.

He has been a consultant to a number of leading financial institutions including the Frank Russell Company, Morgan Stanley, Buchanan Partners, RAB Hedge Funds, Gordon Capital, Matcap, Ketchum Trading and, in the gambling area, to the BC Lotto Corporation, SCA Insurance, Singapore Pools, Canadian Sports Pool, Keeneland Racetrack, and some racetrack syndicates in Hong Kong, Manila, and Australia. His research is in asset-liability management, portfolio theory and practice, security market imperfections, Japanese and Asian financial markets, hedge fund strategies, risk management, sports and lottery investments, and applied stochastic programming. His co-written practitioner paper on the Russell-Yasuda model won second prize in the 1993 Edelman Practice of Management Science Competition. He has been a futures and equity trader and hedge fund and investment manager since 1983.

He has published widely in journals such as *Operations Research, Management Science, Mathematics of OR, Mathematical Programming, American Economic Review, Journal of Economic Perspectives, Journal of Finance, Journal of Economic Dynamics and Control, JFQA, Quantitative Finance, Journal of Portfolio Management*, *and Journal of Banking and Finance* and in many books and special journal issues.

Recent books include *Applications of Stochastic Programming* with S.W. Wallace, SIAM-MPS, 2005, *Stochastic Optimization Models in Finance*, 2nd edition with R.G. Vickson, World Scientific, 2006 and *Handbook of Asset and Liability Modeling*, Volume 1: *Theory and Methodology* (2006) and Volume 2: *Applications and Case Studies* (2007) with S. A. Zenios, North *Holland, Scenarios for Risk Management and Global Investment Strategies* with Rachel Ziemba, Wiley, 2007, *Handbook of Investments: Sports and Lottery Betting Markets*, with Donald Hausch, North Holland, 2008, *Optimizing the Aging, Retirement and Pensions Dilemma* with Marida Bertocchi and Sandra Schwartz and *The Kelly Capital Growth Investment Criterion, 2010,* with legendary hedge fund trader Edward Thorp and Leonard MacLean, *Calendar Anomalies and Arbitrage, The Handbook of Financial Decision Making* (with Leonard MacLean) and *Stochastic Programming* (with Horand Gassman), published by World Scientific in 2012 and 2013. In progress in 2014 are Handbooks on the *Economics of Wine* (with O. Ashenfelter, O. Gergaud and K. Storchmann) and *Futures* (with T. Mallaris).

He is the series editor for North Holland's Handbooks in Finance, World Scientific Handbooks in Financial Economics and Books in Finance, and previously was the CORS editor of INFOR and the department of finance editor of *Management Science*, 1982–1992. He has continued his columns in *Wilmott* and his 2013 book with Rachel Ziemba have the 2007–2013 columns updated with new material published by World Scientific. Ziemba, along with Hausch, wrote the famous *Beat the Racetrack* book (1984) (which was revised into *Dr Z's Beat the Racetrack*, 1987), which presented their place and show betting system, and the *Efficiency of Racetrack Betting Markets* (1994, 2008)—the so-called bible of racetrack syndicates. Their 1986 book *Betting at the Racetrack* extends this efficient/inefficient market approach to simple exotic bets. Ziemba is revising *BATR* into *Exotic Betting at the Racetrack* (World Scientific) which adds Pick3,4,5,6, etc and provides updates to be out in the spring 2014.

Chapter 1
The Theory of Risk, Return, and Performance Measurement

John Guerard

The purpose of this introductory chapter is to trace the role that Jack Treynor played in the development and application of the role of risk in stock price valuation and portfolio performance measurement.[1] In the King's English, the author seeks to acquaint readers with the modern theory of portfolio theory, capital market equilibrium, how institutional investors use modern risk models, and portfolio performance measures. Jack Treynor was at the forefront of these topics some 50 years ago and almost every investment text, many professional journal articles, and the investment practitioner community cite his portfolio performance ratio. One could describe this chapter as "Treynor for the Masses." It is recommended that the reader refer to five reference volumes for this chapter: (1) Treynor on Institutional Investing; (2) William Sharpe, entitled Portfolio Theory and Capital Markets; (3) The Founders of Modern Finance: Their Prize-winning Concepts and 1990 Nobel Lectures; (4) the Christopherson, Carino, and Ferson monograph, entitled Portfolio Performance Measurement and Benchmarking; and the Connor, Goldberg, and Korajczyk monograph, entitled Portfolio Risk Analysis.

Harry Markowitz created a portfolio construction theory in which investors should be compensated with higher returns for bearing higher risk. The Markowitz mean-variance analysis (1952, 1956, 1959) provided the framework for measuring risk, as the portfolio standard deviation. To many people, both academicians and practitioners, the theory of risk measurement and the risk–return trade-off

[1]The author appreciates comments of Professors C.F. Lee and Bernell Stone, Anureet Saxena and Dieter Vandenbussche on earlier drafts of this chapter. Any errors remaining are the sole responsibility of the author.

J. Guerard (✉)
McKinley Capital Mgt, LLC, 3301 C Street, Suite 500, Anchorage, AK 99503, USA
e-mail: JGuerard@McKinleyCapital.com

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_1

begins with Markowitz's seminal volume, entitled Portfolio Selection (1959). In Chap. 4, Markowitz introduced the reader to standard deviation and variances, and covariance and correlations, as measures of dispersion, or risk. Furthermore, Markowitz stated that analysts could and would provide measures of expected returns and variances, but could not reasonably provide estimates of covariance, which could be derived by means of an electronic computer (page 97). Portfolio construction and management, as formulated in Markowitz seeks to identify the efficient frontier, the point at which the portfolio return is maximized for a given level of risk, or equivalently, portfolio risk is minimized for a given level of portfolio return. The portfolio expected return, denoted by $E(R_{\mathrm{P}})$, is calculated by taking the sum of the security weight multiplied by their respective expected return:

$$E\left(R_{\mathrm{P}}\right)=\sum_{i=1}^{N} w_i E\left(R_i\right) \tag{1.1}$$

The portfolio standard deviation is the sum of the weighted securities covariances:

$$\sigma_{\mathrm{p}}^{2}=\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} \tag{1.2}$$

where N is the number of candidate securities, w_i is the weight for security i such that $\sum_{i=1}^{N} w_i = 1$ indicating that the portfolio is fully invested, and $E(R_i)$ is the expected return for security *i*. The Markowitz framework measures risk as the portfolio standard deviation, a measure of dispersion or total risk. One seeks to minimize risk, as measured by the covariance matrix in the Markowitz framework, holding constant expected returns. The decision variables estimated in the Markowitz model are the security weights. The Markowitz model minimized the total risk, or variance, of the portfolio. Investors are compensated for bearing total risk.

Still in Chap. 4, Markowitz suggested that the analyst's team could develop graphical relationships between security returns and the change in the index for each of the other stocks in the analysis (pages 97–98). Thus, we see that Markowitz anticipated the need for quick and efficient calculation of data and the estimation of total and systematic, or market-related, risk. Many readers believe that Markowitz only regarded risk as being measured by variances and covariance, forgetting to read Chap. 9 on the semi-variance, or "down-side risk." Harry recognized the existence of skewness (page 191), and stated quite clearly and correctly that, for a given expected return and variance, investors prefer portfolios with the greatest skewness to the right.

Jack Treynor, in 1962, wrote an unpublished paper that greatly influenced finance, anticipating much of the Capital Asset Pricing Model. This unpublished paper was published in Korajczyk (1999) and reprinted in Treynor (2008) and the paper followed in the steps of Markowitz and Modigliani and Miller (1958, 1963) and sought to provide the groundwork for a theory of valuation that incorporates risk. The unpublished paper is "Toward a Theory of Market Value of Risky Assets." Jack considered an idealized capital market to establish how risk premiums implicit in share prices are related to portfolio decisions of investors without the complexities of taxes and other frictions that may significantly affect share prices in the real world. Treynor listed his assumptions: (1) no taxes; (2) frictions, such as brokerage costs; (3) the individual investor cannot influence the price; (4) investors maximize expected utility, with primary concern for the first and second moments of the distribution of outcomes; (5) investors are risk-averse; (6) a perfect lending market exists; and (7) perfect knowledge, interpreted to mean knowledge of present prices, and what investors know about the future might have a bearing of future investment values. Treynor stated that the expected yield to the investor is a return on his capital at the risk-free lending rate and an expected return for risk-taking.[2] Treynor held that the assumption of risk aversion led to, for a given level of expected returns (performance) in an optimal process, uncertainty is minimized. Jack recognized that an optimal portfolio minimized the portfolio variance subject to the constraint that the expected yield was equal to a constant. The use of Lagrange multipliers led to the result that all efficient combinations have the same ratio of risk premium to standard error. The tangency of the locus of efficient combinations with a utility isoquant determined the expected risk premium for the investor in question. In equilibrium, on page 56 (Treynor, 2008), Jack stated that in his idealized equity market, the risk premium per share of the individual investment is proportional to the covariance of the investment with the total market value of all investments in the market. Thus, in one sentence, Jack Treynor held that the stock risk premium would be proportional to its covariance with the market, and an early statement of the Capital Asset Pricing Model (CAPM). Jack held the econometric problems of estimating the expected performance and covariance were "outside the scope of this study."[3] Portfolio returns were proportional to the covariances among the stocks, not the total risk of the stocks. The covariances among stocks would be recognized in the coming months as stock betas. Jack sought to distinguish risk premium for capital budgeting problems between risks that are assumed independent of market fluctuations (the general level of the market), and those which are assumed not to be. Investments that are risky and independent of market fluctuations are called insurable risks and have a cost of capital equal to the lending rate. Furthermore, the investor holds shares in each equity proportional to the total number of shares available in the market (always positive). As one reads Jack Treynor's 1962 unpublished manuscript, one sees how he builds upon Markowitz (1959) and Tobin (1958), and is contemporary with William (Bill) Sharpe (1963, 1964). Sharpe's (1963) paper, "A Simplified Model for Portfolio Selection" developed the Diagonal Model in which returns of securities are related only through common relationships with an underlying factor; i.e., the stock market as a whole, or Gross National Product (etc.). The portfolio

[2]Treynor (2008), pages 49–51.

[3]Treynor, 2008, p. 57.

variance was the slope of the estimated diagonal model squared times the variance of the market. The covariance of two securities was proportional to the product of their respective slopes of the diagonal model. Sharpe noted that the number of estimates for 100 securities fell from 5150 to only 302, using the Diagonal Market (and for 200 securities, from 2,003,000 to 6002). The time necessary for 100 stock calculations fell from 33 min with quadratic programming codes to 30 s in the diagonal code. Bill Sharpe concluded that the diagonal model "appears to be an excellent choice for the initial practical applications of the Markowitz technique" (page 72).[4] In William Sharpe's (1964) paper, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," Bill stated that capital asset prices must change until a set of prices is "attained for which every asset enters at least one combination lying on the capital market line," the line between risk and expected return, with a intercept of the risk-free rate. Sharpe then demonstrated a consistent relationship between individual asset's expected returns and systematic risk, the responsiveness of an asset's rate of return to the level of economic activity.[5] Prices adjust until there is a linear relationship between the magnitude of the responsiveness and expected returns (page 93).

The Treynor (1999), Sharpe (1964), Lintner (1965a), and Mossin (1966) developments of the Capital Asset Pricing Model (CAPM) held that investors are compensated for bearing not total risk, but rather market risk, or systematic risk, as measured by a stock's beta; investors are not compensated for bearing stock-specific risk, which can be diversified away in a portfolio context. A stock's beta is the slope of the stock's return regressed against the market's return. Modern capital theory has evolved from one beta, representing market risk, to multi-factor risk models (MFMs), many with four or more betas. Stone (1970) presented a framework for asset selection and capital market equilibrium.

Investment managers seeking the maximum return for a given level of risk create portfolios using many sets of models, based both on historical and expectation data. In this introductory chapter, we briefly trace the evolution of the estimated models of risk and show how risk models enhance portfolio construction, management, and evaluation.

Let us briefly provide a roadmap of the remaining chapter for the reader. First, the reader is introduced to Capital Market Theory and the relationship between the Capital Market Line and the Security Market Line. Second, we discuss the historical developments and approaches to estimation of MFMs. We particularly emphasize the BARRA MFM, as it is the risk model most widely used by asset managers.

[4] The Founders of Modern Finance: Their Prize-Winning Concepts and 1990 Nobel Lectures, The Research Foundation of the Institute of Chartered Financial Analysts. 1991.

[5] Bill Sharpe noted, in his footnote 7, on page 77, that upon completion of his paper, that he learned that Mr. Jack L. Treynor, of Arthur D. Little, Inc., had independently developed a model similar in many respects, to his. Jack's excellent work, he noted is, at present, unpublished. In summary, Treynor's portfolio returns were proportional to the covariances among the stocks and market index, not the total risk of the stocks. The covariances among stocks and the market index would be recognized in the coming months as stock betas by Bill Sharpe.

Third, we introduce the reader to an alternative risk model vendor, Axioma, which estimates statistical and fundamental risk models that minimizes the tracking error of portfolios for asset managers. Fourth, we review the Lee, Lee, and Liu (2010) findings on mutual fund performance using the Treynor (1965) and Treynor–Mazuy (1966) indexes.

1.1 Capital Market Equilibrium

Let us review portfolio theory and capital market equilibrium. Sharpe (1970) discusses capital market theory in which investors purchase or sell stocks based on beliefs, or predictions, of expected returns, standard deviations of returns, and correlation coefficients of returns. Indeed, all investors have identical expectations of the predictions, known as homogeneous beliefs. Investors seek to maximize return while minimizing risk. Investors may lend and borrow as much as they desire at the pure rate of interest, also known as the risk-free rate of interest. Capital market theory holds that once equilibrium is established, then it is maintained. In equilibrium, there is no pressure to change. Sharpe states that capital market theory asks about relationship between expected returns and risk for (1) portfolios and (2) securities. The implicit question concerns the appropriate measure of risk for (1) a portfolio and (2) a security. The optimal combination of risky securities is the market portfolio, which is the percentage of market value of the security as compared to the total market value of risky securities. The market portfolio includes only risky assets, and the actual return on the market portfolio is the weighted average of the expected returns of the risky securities. A linear line passes from the risk-free interest rate through the market portfolio on a return-risk graph. An individual preference curve, known as a indifference curve, determines where along this line, known as the Capital Market Line, the investor seeks to invest. A conservative investor might seek to lend money at the risk-free rate and invest the remaining funds in the market portfolio. An aggressive investor might borrow money at the risk-free rate and invest more funds than his or her initial endowment in the market portfolio. To increase expected return along the Capital Market Line, the investor must accept more risk. Conversely, to reduce risk, an investor must give up expected return. Sharpe (1970) refers to the slope of the Capital Market Line as the price of risk reduction (in terms of decreased expected return). All efficient portfolios must lie along the capital market line where:

$$E\left(R_{\mathrm{p}}\right) = R_{\mathrm{F}} + r_{\mathrm{e}}\sigma_{\mathrm{p}} \tag{1.3}$$

Where $E(R_{\mathrm{P}}) =$ expected portfolio return,
$r_{\mathrm{e}} =$ price of risk reduction for efficient portfolios,
$R_{\mathrm{F}} =$ pure (risk-free) rate of interest,
and $\sigma_{\mathrm{p}} =$ portfolio standard deviation.

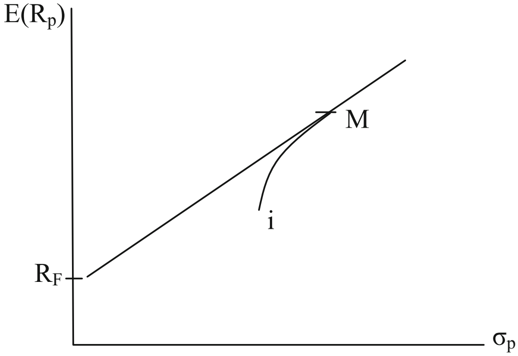


Fig. 1.1 The Security Market Line

The Capital Market Line, Eq. (1.1), summarizes the simple (linear) relationship between expected return and risk of efficient portfolios. However, such a relationship does not hold for inefficient portfolios and individual securities.

Sharpe (1970) presents a very reader-friendly derivation of the Security Market Line (SML). Sharpe assumed that total funds were divided between the market portfolio, M, and security i. The investor is fully invested; hence

$$X_M + x_i = 1$$

The expected return of the portfolio is:

$$E\left(R_{\mathrm{p}}\right) = x_i E\left(R_i\right) + x_M E\left(R_M\right) \tag{1.4}$$

and the corresponding variance of the portfolio, z, is:

$$\sigma_{\mathrm{p}} = {x_i}^2 {\sigma_i}^2 + x_M^2\,\sigma_M^2 + 2x_i x_M \rho_{iM}\sigma_i\sigma_M \tag{1.5}$$

We saw these equations in the previous chapter and see no reason for numbering them. Assume that $E(R_i)$ is less than $E(R_M)$, as is its standard deviation. Sharpe presents this graphically as Fig. 1.1.

We know that the slope of the curve connecting the security i and market portfolio M depends upon the relative weights, x_i and x_M, and the correlation coefficient, ρ_{iM}, between the security i and market portfolio M.

Rewrite $x_M = 1{-}x_i$ and $\sigma_{iM} = \rho_{iM}\ \sigma_i\ \sigma_M$.

$$\sigma_{\mathrm{p}} = \sqrt{x_i^2\sigma_i^2 + (1 - x_i)^2\sigma_M^2 + 2x_i\,(1 - x_i)\,\sigma_{iM}}$$

To find an optimal weight:

$$\frac{\partial\sigma_{\mathrm{p}}}{\partial x_i} = \frac{x_i\left(\sigma_i^2 + \sigma_M^2 - 2\sigma_{iM}\right) + \sigma_{iM} - \sigma_M^2}{\sigma_z} \tag{1.6}$$

In order to minimize portfolio risk, one takes the derivative of total risk relative to the portfolio decision variable, the security weight.

$$E\left(R_{\mathrm{p}}\right) = x_i E\left(R_i\right) + \left(1 - x_i\right) E\left(R_M\right)$$

$$\frac{\partial E_{\mathrm{p}}}{\partial x_i} = E\left(R_i\right) - E\left(R_M\right) \tag{1.7}$$

One uses Eq. (1.2) and the chain rule to give:

$$\frac{\partial E_{\mathrm{p}}}{\partial \sigma_{\mathrm{p}}} = \frac{\partial E_{\mathrm{p}}/\partial x_i}{\partial \sigma_{\mathrm{p}}/\partial x_i} = \frac{E\left(R_i\right) - E\left(R_M\right)}{\frac{x_i\left(\sigma_i^2+\sigma_M^2-\sigma_{iM}\right)+\sigma_{iM}-\sigma_M^2}{\sigma_z}} \tag{1.8}$$

Equation (1.6) shows the trade-off between expected returns and portfolio standard deviations. At point M, the market portfolio, $x_i = 0$ and $\sigma_{\mathrm{p}} = \sigma_M$, thus:

$$\frac{\partial E_{\mathrm{p}}}{\partial \sigma_{\mathrm{p}}} /_{x_i=0} = \frac{E\left(R_i\right) - E\left(R_M\right)}{\left(\sigma_{iM} - \sigma_M^2\right) / \sigma_M} = \frac{\left[E\left(R_i\right) - E\left(R_M\right)\right] \sigma_M}{\sigma_{iM} - \sigma_M^2} \tag{1.9}$$

In equilibrium, curve *iM* becomes tangent to the capital Market Line. The investor must be compensated for bearing risk.

The curve *iM* and slope of the Capital Market Line must be equal. The trade-off of expected return and risk for the portfolio must be equal to the capital market trade-off.

Thus

$$\frac{\left[E\left(R_i\right) - E\left(R_M\right)\right] \sigma_M}{\sigma_{iM} - \sigma_M^2} = \frac{E\left(R_M\right) - R_{\mathrm{F}}}{\sigma_M} \tag{1.10}$$

or

$$E\left(R_i\right) - R_{\mathrm{F}} = \left[\frac{E\left(R_M\right) - R_{\mathrm{F}}}{\sigma_M^2}\right] \sigma_{iM} \tag{1.11}$$

and

$$E\left(R_i\right) = R_{\mathrm{F}} + \left[E\left(R_M\right) - R_{\mathrm{F}}\right] \frac{\sigma_{iM}}{\sigma_M^2} \tag{1.12}$$

Sharpe (1970) discusses the stock beta as the slope of the firm's characteristic line, the volatility of the security's return relative to changes in the market return.

The CAPM holds that the return to a security is a function of the security's beta.

$$R_{jt} = R_{\mathrm{F}} + \beta_j \left[E\left(R_{Mt}\right) – R_{\mathrm{F}}\right] + e_{jt} \tag{1.13}$$

where R_{jt} = expected security j return at time t;
$E\ (R_{Mt})$ = expected return on the market at time t;
R_F = risk-free rate;
β_j = security beta, a random regression coefficient; and
e_j = randomly distributed error term.

Let us examine the Capital Asset Pricing Model beta, its measure of systematic risk, from the Capital Market Line equilibrium condition, in an alternative formulation.

$$\beta_j = \frac{\text{Cov}\left(R_j, R_M\right)}{\text{Var}\left(R_M\right)} \tag{1.14}$$

$$\begin{aligned} E\left(R_j\right) &= R_F + \left[\frac{E(R_M) - R_F}{\sigma_M{}^2}\right] \text{Cov}\ \left(R_j, R_M\right) \\ &= R_F + \left[E\left(R_M\right) - R_F\right] \frac{\text{Cov}\left(R_j, R_M\right)}{\text{Var}\left(R_M\right)} \end{aligned}$$

$$E\left(R_j\right) = R_F + \left[E\left(R_M\right) - R_F\right] \beta_j \tag{1.15}$$

Equation (1.13) defines the Security Market Line (SML), which describes the linear relationship between the security's return and its systematic risk, as measured by beta.

1.2 The Barra Model: The Primary Institutional Risk Model

No sooner was the CAPM developed then its estimations became of concern. Black, Jensen, and Scholes (1972) found second-pass regressions with smaller than expected slopes and higher than expected intercepts, implying investor were notes risk-averse as they "should be." Elton and gruber (1970), Farrell (1974, 1997), Stone (1974), Rosenberg and McKibben (1973), and Rosenberg (1974) proposed and estimated alternative forms of multi-factor models. The majority of institutional investors have been introduced to risk through the Barra risk model. In 1975, Barr Rosenberg and his associates introduced the Barra US Equity Model, often denoted USE1. We spend a great deal of time on the Barra USE1 and USE3 models because 70 of the 100 largest investment managers use the BARRA US Equity Risk Models.[6] The Barra USE1 Model predicted risk and risk was measured by the company's fundamentals, as discussed in Rosenberg and Rudd (1977), Rudd and Clasing (1982).[7] The US Equity Models have been updated and enhanced; USE2 in 1985, USE3 in 1997, and USE4 in 2011. The first Barra Global Equity

[6] According to BARRA online advertisements.

[7] There are several definitive treatments of the Barra system. Rosenberg and Marathe (1979) was a seminal test of capital asset pricing. Rudd and Clasing (1982), Grinold and Kahn (2000), and Menchero, Morozov, and Shepard (2010) are some of the most cited Barra model publications.

Risk Model (GEM) was introduced in 1989. The model was estimated via monthly cross-sectional regressions using countries, industries, and styles as explanatory factors, as described by Grinold, Rudd, and Stefek (1989). GEM2 was introduced in 2008 and published in Menchero et al. (2010). GEM2 incorporated several advances over the previous model, such as improved estimation techniques, higher-frequency observations, and the introduction of the World factor to place countries and industries on an equal footing. GEM3 was introduced in 2011. Given that Barra, now (Morgan Stanley Capital International) MSCI Barra is the most successful commercial risk model and most widely used risk model, it seems appropriate for us to discuss its creation and structure. The reader is referred to Rudd and Clasing (1982) for a complete treatment of the Barra model development. Much of this section drawn heavily from Rudd and Clasing.

Barr Rosenberg and Walt McKibben (1973) estimated the determinants of security betas and standard deviations. This estimation formed the basis of the Rosenberg extra-market component study (1974), in which security specific risk could be modeled as a function of financial descriptors, or known financial characteristics of the firm. The Rosenberg and Marathe (1979) paper developed the econometric methodology of investment returns analysis. Capital market equilibrium linked both first (mean) and second (variance) moments. Rosenberg and Marathe discussed econometric estimation techniques for second moments.[8] The total excess return for a multiple factor model, referred to as the MFM, in the Rosenberg methodology for security j, at time t, dropping the subscript t for time, may be written:

$$E\left(R_j\right) = \sum_{k=1}^{K} \beta_{jk}\tilde{f}_k + \tilde{e}_j \qquad (1.16)$$

The non-factor, or asset-specific, return on security j, is the residual risk of the security, after removing the estimated impacts of the K factors. The term, f, is the rate of return on factor k. A single factor model, in which the market return is the only estimated factor, is obviously the basis of the Capital Asset Pricing Model. Accurate characterization of portfolio risk requires an accurate estimate of the covariance matrix of security returns. A relatively simple way to estimate this covariance matrix is to use the history of security returns to compute each variance, covariance and security beta. The use of beta, the covariance of security and market index returns, is one method of estimating a reasonable cost of equity funds for firms. However, the approximation obtained from simple models may not yield the best possible cost of equity. The simple, single index beta estimation approach suffers from two major drawbacks:

- Estimating a covariance matrix for the Russell 3000 stocks requires a great deal of data;

[8] Barr Rosenberg and several coauthors put forth some 12–15 papers during the 1973–1979 time period that created the intellectual basis of the Barra system. The Rosenberg publications were both academic and practitioner journals.

- It is subject to estimation error. Thus, one might expect a higher correlation between DuPont and Dow than between DuPont and IBM, given that DuPont and Dow are both chemical firms.

Taking this further, one can argue that firms with similar characteristics, such as firms in their line of business, should have returns that behave similarly. For example, DuPont and IBM will all have a common component in their returns because they would all be affected by news that affects the stock market, measured by their respective betas. The degree to which each of the three stocks responds to this stock market component depends on the sensitivity of each stock to the stock market component.

Additionally, one would expect DuPont and Dow to respond to news effecting the chemical industry, whereas IBM and Dell would respond to news effecting the Computer industry. The effects of such news may be captured by the average returns of stocks in the chemical industry and the computer industry. One can account for industry effects in the following representation for returns:

$$\begin{aligned}\tilde{r}_{\text{DD}} = {} & E\left[\tilde{r}_{\text{DD}}\right] + \beta \cdot \left[\tilde{r}_M - E\left[\tilde{r}_M\right]\right] \\ & + 1 \cdot \left[\tilde{r}_{\text{CHEMICAL}} - E\left[\tilde{r}_{\text{CHEMICAL}}\right]\right] + 0 \cdot \left[\tilde{r}_{\text{C}} - E\left[\tilde{r}_{\text{DD}}\right]\right] + \mu_{\text{P}}\end{aligned} \tag{1.17}$$

where:
$\tilde{r}_{\text{DD}}$ = DD's realized return,
$\tilde{r}_M$ = the realized average stock market return,
$\tilde{r}_{\text{CHEMICAL}}$ = realized average return to chemical stocks,
$\tilde{r}_{\text{C}}$ = the realized average return to computer stocks,
$E[.]$ = expectations
β_{DD} = DD's sensitivity to stock market returns, and
μ_{DD} = the effect of DD specific news on DD returns.

This equation simply states that DD's realized return consists of an expected component and an unexpected component. The unexpected component depends on any unexpected events that affect stock returns in general $[\tilde{r}_M - E[\tilde{r}_M]]$, any unexpected events that affect the chemical industry $[\tilde{r}_{\text{CHEMICAL}} - E[\tilde{r}_{\text{CHEMICAL}}]]$, and any unexpected events that affect DD alone (μ_{DD}). Thus, the sources of variation in DD's stock returns, are variations in stock returns, in general, variations in chemical industry returns, and any variations that are specific to DD. Moreover, DD and Dow returns are likely to move together because both are exposed to stock market risk and chemical industry risk. DD, IBM, and D, Dominion Resources, on the other hand, are likely to move together to a lesser degree because the only common component in their returns is the market return.

Investors look at the variance of their total portfolios to provide a comprehensive assessment of risk. To calculate the variance of a portfolio, one needs to calculate the covariances of all the constituent components. Without the framework of a multiple-factor model, estimating the covariance of each asset with every other asset is computationally burdensome and subject to significant estimation errors. The Barra MFM simplifies these calculations dramatically, replacing individual company profiles with categories defined by common characteristics (factors). The

$$\tilde{r} = X\tilde{f} + \tilde{u}$$

where $\tilde{r}$ = vector of excess returns,

X = exposure matrix,

$\tilde{f}$ = vector of factor returns, and

$\tilde{u}$ = vector of specific returns.

$$\begin{bmatrix} \tilde{r}(1) \\ \tilde{r}(2) \\ \vdots \\ \tilde{r}(N) \end{bmatrix} = \begin{bmatrix} X(1,1) & X(1,2) & \dots & X(1,K) \\ X(2,1) & X(2,2) & \dots & X(2,K) \\ \vdots & \vdots & & \vdots \\ X(N,1) & X(N,2) & \dots & X(N,K) \end{bmatrix} \begin{bmatrix} \tilde{f}(1) \\ \tilde{f}(2) \\ \vdots \\ \tilde{f}(K) \end{bmatrix} + \begin{bmatrix} \tilde{u}(1) \\ \tilde{u}(2) \\ \vdots \\ \tilde{u}(N) \end{bmatrix}$$

Fig. 1.2 The Barra MFM Risk Model

specific risk is assumed to be uncorrelated among the assets and only the factor variances and covariances are calculated during model estimation (Fig. 1.2).

The multiple-factor risk model, significantly reduces the number of calculations inherent in covariance analyses. For example, in the US Equity Model (USE3), 65 factors capture the risk characteristics of equities. This reduces the number of covariance and variance calculations; moreover, since there are fewer parameters to determine, they can be estimated with greater precision. The BARRA risk management system begins with the MFM equation:

$$\tilde{r}_i = X\tilde{f} + \tilde{u} \tag{1.18}$$

where:
$\tilde{r}_i$ = excess return on asset i,
X = exposure coefficient on the factor,
$\tilde{f}$ = factor return, and
$\tilde{u}$ = specific return.

Substituting this relation in the basic equation, we find that:

$$\text{Risk} = \text{Var}\,(\tilde{r}_i) \tag{1.19}$$

$$= \text{Var}\left(X\tilde{f} + \tilde{u}\right) \tag{1.20}$$

Using the matrix algebra formula for variance, the risk equation becomes:

$$\text{Risk} = XFX^{\text{T}} + \Delta \tag{1.21}$$

where:

X = exposure matrix of companies upon factors,

F = covariance matrix of factors,

X^T = transpose of X matrix, and

Δ = diagonal matrix of specific risk variances.

This is the basic equation that defines the matrix calculations used in risk analysis in the BARRA equity models.[9] Investment managers seek to maximize portfolio return for a given level of risk. For many managers, risk is measured by the BARRA risk model.[10]

The most frequent approach to predict risk is to use historical price behavior in the estimation of beta. Beta was defined as the sensitivity of the expected excess rate of return on the stock to the expected excess rate of return on the market portfolio. A major assumption has to be made to enable average (realized) rates of return to be used in place of expected rates of return, which, in turn, permits one to use the slope of regression line (estimated from realized data) to form the basis for a prediction of beta.

If this assumption, which essentially states that the future is going to be similar to the "average past," is made, then the estimation of historical beta proceeds as follows. When more data points are used, the accuracy of the estimation procedure is improved, provided the relationship being estimated does not change. Usually the relationship does change; therefore, a small number of most recent data points is preferred so that dated information will not obscure the current relationship. It is usually accepted that a happy medium is achieved by using 60 monthly returns.[11] The security series is then regressed against the market portfolio series. This provides an estimate of beta (which is equivalent to the slope of the characteristic line) and the residual variance.

It can be shown that if the regression equation is properly specified and certain other conditions are fulfilled, then the beta obtained is an optimal estimate (actually, minimum-variance, unbiased) of the true historical beta averaged over past periods. However, this does not imply that the historical beta is a good predictor of future beta. For instance, one defect is that random events impacting the firm in the

[9]Markowitz discusses the MFM formulation in his second monograph, Mean-Variance Analysis in Portfolio Choice and Capital Markets, New Hope, PA: Frank J. Fabozzi Associates, 2000), Chapter 3, pp. 45–47. The Markowitz (1987, 2000) Mean-Variance Analysis volume requires great patience and thought on the part of the reader, as noted by Bill Sharpe in his foreword to the 2000 edition.

[10]Jose Menchero and his colleagues at BARRA authored the "Global Equity Risk Modeling" article in the Markowitz volume, estimated an eight-risk global index model in the spirit of the Rosenberg USE3 model, see Guerard (2010).

[11]We have glossed over a number of econometric subtleties in these few sentences. Those readers who wish to learn more about these estimation difficulties are directed toward the following articles and the references contained there: Merton Miller and Myron Scholes, "Rates of Return in Relation to Risk: A Reexamination of Recent Findings," in *Studies in The Theory of Capital Markets*, ed. Michael Jensen (New York: Praeger Publishers, 1972), pp. 47–48 and Eugene F. Fama, Foundations of Finance (New York: Basic Books, 1976), Chapter 4.

past may have coincided with market movements purely by chance, causing the estimated value to differ from the true value. Thus, the beta obtained by this method is an estimate of the true historical beta obscured by measurement error. Rudd and Clasing (1982) discussed beta prediction with respect to the use of historic price information. Three possible prediction methods for beta were suggested. These are:

1. *Naïve*: $\widehat{\beta}_j = 1.0$ for all securities (i.e., every security has the average beta).
2. *Historical*: $\widehat{\beta}_j = \mathrm{H}\widehat{\beta}_j$, the historical beta obtained as the coefficient form an ordinary least squares regression.
3. *Bayesian adjusted beta*: $\widehat{\beta}_j = 1.0 + \mathrm{BA}\left(H\widehat{\beta}_j - 1\right)$, where the historical betas are adjusted toward the mean value of 1.0.

In each case, the prediction of residual risk is obtained by subtracting the systematic variance $\left(\widehat{\beta}_j^2 V_M\right)$ from the total variance of the security. The residual variance is obtained directly from the regression.

However, relying simply upon historical price data is unduly restricting in that there are excellent sources of information which may help in improving the prediction of risk. For instance, most analysts would agree that fundamental information is useful in understanding a company's prospects.

The historical beta estimate will be an unbiased predictor of the future value of beta, provided that the expected change between the true value of beta averaged over the past periods and its value in the future is zero. If this expected change is not zero, then the historical beta estimate will be misleading (biased). Thus, if historical betas are used as a prediction of beta, there is an implicit assumption that the future will be similar to the past. Is this assumption reasonable? The investment environment changes so rapidly that it would appear imprudent to use averages of historical (5-year) price data as predictions of the future.

The empirical evidence regarding the construction of the Barra risk models comes from several well-known publications by Barr Rosenberg. Barr Rosenberg and Walt McKibben (1973) estimated the determinants of security betas and standard deviations. This estimation formed the basis of the Rosenberg extra-market component study (1974), in which security specific risk could be modeled as a function of financial descriptors, or known financial characteristics of the firm. Rosenberg and McKibben found that the financial characteristics that were statistically associated with beta during the 1954–1970 period were:

1. Latest annual proportional change in earnings per share;
2. Liquidity, as measured by die quick ratio;
3. Leverage, as measured by die senior debt-to-total assets ratio;
4. Growth, as measured by the growth in earnings per share;
5. Book-to-Price ratio;
6. Historic beta;
7. Logarithm of stock price;
8. Standard deviation of earnings per share growth;
9. Gross plant per dollar of total assets;
10. Share turnover.

Rosenberg and McKibben used 32 variables and a 578-firm sample to estimate the determinants if betas and standard deviations. For betas, Rosenberg and McKibben found that the positive and statistically significant determinants of beta were the standard deviation of EPS growth, share turnover, the price-to-book multiple, and the historic beta.[12] The statistically significant determinants of the security systematic risk became the basis of the Barra E1 Model risk indexes.

The Barra USE1 Model predicted risk, which required the evaluation of the firm's response to economic events, which were measured by the company's fundamentals,

[12]When an analyst forms a judgment on the likely performance of a company, many sources of information can be synthesized. For instance, an indication of future risk can be found in the balance sheet and the income statement; an idea as to the growth of the company can be found from trends in variables measuring the company's position; the normal business risk of the company can be determined by the historical variability of the income statement; and so on. The approach that Rosenberg and Marathe take is conceptually similar to such an analysis since they attempt to include all sources of relevant information. This set of data includes historical technical and fundamental accounting data. The resulting information is then used to produce, by regression methods, the fundamental predictions of beta, specific risk, and the exposure to the common factors.

The fundamental prediction method of Barra starts by describing the company, see Rudd and Clasing (1982). The Barra USE1 Model estimated "descriptors," which are ratios that describe the fundamental condition of the company. These descriptors are grouped into six categories to indicate distinct sources of risk. In each case, the category is named so that a higher value is indicative of greater risk.

1. *Market variability*. This category is designed to capture the company as perceived by the market. If the market was completely efficient, then all information on the state of the company would be reflected in the stock price. Here the historical prices and other market variables are used in an attempt to reconstruct the state of the company. The descriptors include historical measures of beta and residual risk, nonlinear functions of them, and various liquidity measures.
2. *Earnings variability*. This category refers to the unpredictable variation in earnings over time, so descriptors such as the variability of earnings per share and the variability of cash flow are included.
3. *Low valuation and unsuccess*. How successful has the company been, and how is it valued by the market? If investors are optimistic about future prospects and the company has been successful in the past (measured by a low book-to-price ratio and growth in per share earnings), then the implication is that the firm is sound and that future risk is likely to be lower. Conversely, an unsuccessful and lowly valued company is more risky.
4. *Immaturity and smallness*. A small, young firm is likely to be more risky than a large, mature firm. This group of descriptors attempts to capture this difference.
5. *Growth orientation*. To the extent that a company attempts to provide returns to stockholders by an aggressive growth strategy requiring the initiation of new projects with uncertain cash flows rather than the more stable cash flows of existing operations, the company is likely to be more risky. Thus, the growth in total assets, payout and dividend policy, and the earnings/price ratio are used to capture the growth characteristics of the company.
6. *Financial risk*. The more highly levered the financial structure, the greater is the risk to common stockholders. This risk is captured by measures of leverage and debt to total assets.

Finally industry in which the company operates is another important source of information. Certain industries, simply because of the nature of their business, are exposed to greater (or lesser) levels of risk (e.g., compare airlines versus gold stocks). Rosenberg and Marathe used indicator (dummy) variables for 39 industry groups as the method of introducing industry effects.

as discussed in Rudd and Clasing (1982). There are three major steps. First, for the time period during which the model is to be fitted, obtain common stock returns and company annual reports (for instance, from the COMPUSTAT data base).[13] In order to make comparisons across firms meaningful, the descriptors must be normalized so that there is a common origin and unit of measurement.

The normalization takes the following form. First, the "raw" descriptor Values for each company are computed. Next, the capitalization weighted value of each descriptor for all the securities in the S&P 500 is computed and then subtracted from each raw descriptor. The transformed descriptors now have the property that the capitalization weighted value for the S&P 500 stocks is zero. Furthermore, the standard deviation of each descriptor is calculated within a universe of large companies (defined as having a capitalization of $50 million or more). The descriptor is now further transformed by setting the value + 1 to be one standard deviation above the S&P 500 mean (i.e., one unit of length corresponds to one standard deviation). Rudd and Clasing (1982) write:

$$\mathrm{ND} = (\mathrm{RD{-}RD}\ [\mathrm{S\&P}])\,/\mathrm{STDEV}\,[\mathrm{RD}]\,, \tag{1.22}$$

where

ND = the normalized descriptor value;
RD = the raw descriptor value as computer from the data;
RD[S&P] = the raw descriptor value for the (capitalization-weighted) S&P 500;

and

STDEV[RD] = the standard deviation of the raw descriptor value calculated from the universe of large companies.

At this stage each company is identified by a series of descriptors which indicate its fundamental position. If a descriptor value is zero, then the company is "typical" of the S&P 500 (for this characteristic) because the S&P 500 and the company both have the same raw value. Conversely, if the descriptor value is nonzero, then the company is atypical of the S&P 500, and this information may he used to adjust the prior prediction in order to obtain a better posterior prediction of risk.

In the second step, one groups the monthly data by quarters, and assemble the descriptors of each company as they would have appeared at the beginning of the quarter. The prediction rule is then fitted by linear regression which relates each monthly stock return in that quarter to the previously computed descriptors. These adjustments are combined as follows. Initially, in the absence of any fundamental information, the beta is set equal to its historical value. Then each descriptor is examined in turn, and if it is atypical, the corresponding adjustment to beta is made. For example, if two companies with the same historical beta are identical except that they have very different capitalizations, then one adjusts the risk of the

[13] The COMPUSTAT data base is one of the data bases collected by Investors Management Sciences, Inc., a subsidiary of Standard & Poor's Corporation.

large-capitalization company downward, relative to that of the small-capitalization company, because large companies typically have less risk than small companies.[14] The fundamental knowledge of additional information improves the prediction of risk. The econometric prediction rule is similar; the prediction is obtained by adding the adjustments for all descriptors, in addition to the industry effect, to the historical beta estimate. The prediction rule for the beta of security *i*, in a given month, can be written as follows:

$$\widehat{\beta}_i = \widehat{b}_0 + \widehat{b}_1 d_{1i} + \cdots + \widehat{b}_J d_{ji} \tag{1.23}$$

where

$\widehat{\beta}_i =$ the predicted beta;
$\widehat{b}_J =$ the estimated response coefficients in the prediction rule;
$d_{ji} =$ the normalized descriptor values for security i; and
$J =$ the total number of descriptors.

In this prediction rule we can think of the first descriptor, d_{1i}, as the historical beta, H$\widehat{\beta}$. Thus, if only the first descriptor is used, the prediction rule is similar to the specification of the Bayesian adjustment, Eq. (1.23). In this case, the linear regression provides estimates for $\widehat{b}_0$ and $\widehat{b}_1$, which indicate the optimal adjustment to historical beta for predictive purposes. Other descriptors in addition to historical beta are employed and appear in the prediction rule as d_{2i}.

If the company is completely typical of market (i.e., the descriptors other than historical beta are all zero), then there is no further adjustment to the Bayesian-adjusted historical beta. If the company is atypical, then not all the descriptors (other than historical beta) will be zero. The prediction rule, Eq. (1.23), shows that the predicted beta is found by adding the adjustment $\widehat{b}_2$ to the Bayesian-adjusted historical beta. In general, the total adjustment is the weighted sum of the coefficients in the prediction rule, where the weights are the normalized descriptor values which indicate the company's degree of deviance from the typical company. The Barra risk model estimates the company's exposure to each of the common factors and the prediction of the residual risk components. The first task is to form summary measures or indices of risk to describe all aspects of the company's investment risk. These are obtained by forming the weighted average of the descriptor values in each of the six categories introduced above, where the weights are the estimated coefficients from the prediction rule, Eq. (1.23), for systematic or residual risk. This provides six summary measures of risk, the risk indices, for each company. Again, these indices are normalized so that the S&P 500 has a value of zero on each index and a value of one corresponds to one standard deviation among all companies with capitalization of $50 million or more.

[14] Rosenberg and Marathe (1979) used generalized least squares (GLS) to scale firms in their universe of 315,000 observations to produce constant variances.

The prediction of residual risk is now found by performing a regression on the cross section of all security residual returns as the dependent variable where the independent variables are the risk indices.[15] The form of the regression, in a given month, is shown in Eq. (1.24):

$$r_i - \widehat{\beta}_i r_M = c_1 \mathrm{RI}_{1i} + \cdots + c_6 \mathrm{RI}_{6i} + c_7 \mathrm{IND}_{1i} + \cdots + c_{45} \mathrm{IND}_{39,i} + u_i \tag{1.24}$$

where

$r_i =$ the excess return on security i,
$\widehat{\beta}_i =$ the predicted beta, from Eq. (1.9); and
$r_M =$ the excess return on the market portfolio,

so that $r_i - \widehat{\beta}_i r_M$ is the residual return on security i; $\mathrm{RI}_{1i}, \ldots, \mathrm{RI}_{6i}$, are the six risk indices for security i, $\mathrm{IND}_{1i}, \ldots, \mathrm{IND}_{39,i}$ are the dummy variables for the 39 industry groups; u_i, is the specific return for security i; and $c_1, \ldots, c_{45}$ are the 45 coefficients to be estimated.[16]

The entire risk of the security arises from two sources: the systematic or factor risk ($b_j^2 \mathrm{Var}[f]$), and the nonfactor risk (σ_j^2). In this case, however, the nonfactor risk is completely specific risk since no risk arises from interactions with other securities. In other words, under these assumptions the single factor, f, is responsible for the only commonality among security returns; thus, the random return component that is not related to the factor must be specific to the individual security, j.

If we form a portfolio, P, with weights $h_{\mathrm{P}1}, h_{\mathrm{P}2}, \ldots, h_{\mathrm{P}N}$, from N stocks, then the random excess return on the portfolio is given by:

$$R_{\mathrm{P}} = \Sigma h_{\mathrm{P}j} r_j = \Sigma h_{\mathrm{P}j} b_j f + \Sigma h_{\mathrm{P}j} u_j = b_{\mathrm{P}} f + \Sigma h_{\mathrm{P}j} u_j, \tag{1.25}$$

where $b_{\mathrm{P}} = \Sigma h_{\mathrm{P}j} r_j$. The mean return and variance are:

$$E\left[r_{\mathrm{P}}\right] = a_{\mathrm{P}} + b_{\mathrm{P}} E\left[f\right],$$

where $a_{\mathrm{P}} = \Sigma h_{\mathrm{P}j} a_j$, and

$$\mathrm{Var}\left[r_{\mathrm{P}}\right] = b_j^2\ \mathrm{Var}\left[f\right] + \Sigma h_{\mathrm{P}}^2 \sigma_j^2 \tag{1.26}$$

[15] See Barr Rosenberg and Vinay Marathe, "Common Factors in Security Returns: Microeconomic Determinants and Macroeconomic Correlates," *Proceedings of the Seminar on the Analysis of Security Prices*, University of Chicago, May 1976, pp. 61–115.

[16] The result from the cross-sectional regression Eq. (1.23) is the specific return and specific risk on the security, together with the 45 coefficients. These estimated coefficients represent the returns that can be attributed to the factors in the month of the analysis.

where we have made use of the fact that the security-specific risk is *specific*, i.e., independent across securities and independent of the factor return. The regression coefficient of an individual stock's rate of return onto the market, or beta, is given by:

$$\begin{aligned}\beta_j &= \mathrm{Cov}\left[r_j,\ r_M\right]/\mathrm{Var}\left[r_M\right]\\ &= \mathrm{Cov}\left[b_i f + u_j,\ f + \Sigma h_{Mk} u_k\right]/\mathrm{Var}\left[r_M\right]\\ &= \left(b_j \mathrm{Var}\left[f\right] + \sigma_j^2 h_{Mj}\right)/\mathrm{Var}\left[r_M\right]\\ &= \left(b_j \mathrm{Var}\left[f\right] + h_{Mj}\sigma_j^2\right)/\mathrm{Var}\left[f\right] + \Sigma h_{Mj}^2 \sigma_j^2\Big),\end{aligned} \tag{1.27}$$

so that:

$$\beta_{\mathrm{P}} = \left(b_{\mathrm{P}} \mathrm{Var}\left[f\right] + \Sigma h_{Mj} h_{\mathrm{P}j} \sigma_j^2\right)/\mathrm{Var}\left[f\right] + \Sigma h_{Mj}^2 \sigma_j^2\Big).$$ [17]

The domestic BARRA E3 (USE3, or sometimes denoted US-E3) model, with some 15 years of research and evolution, uses 13 sources of factor, or systematic, exposures. The sources of extra-market factor exposures are volatility, momentum, size, size non-linearity, trading activity, growth, earnings yield, value, earnings variation, leverage, currency sensitivity, dividend yield, and non-estimation universe.[18]In November 2011, Barra introduced USE4(L), the Barra US Equity Model

[17]Rudd and Clasing (1982) note that the regression coefficient on the market and the regression coefficient on the factor (i.e., b_j and β_j, and b_{P} and β_{P}) are close but not identical. In other words, for well-diversified portfolios, the majority of institutional portfolios, one can approximate the portfolio beta by its regression coefficient on the factor, and vice versa, that is, $\beta_{\mathrm{P}} \cong b_{\mathrm{P}}$. In a multiple factor model, the security beta is a weighted average of the factor betas and the beta of the specific return of the security, where the weights are simply the factor loadings for the jth security.

[18]The USE3 extra-market factor are composed of:

1. **Volatility** is composed of variables including the historic beta, the daily standard deviation, the logarithm of the stock price, the range of the stock return relative to the risk-free rate, the options pricing model standard deviation, and the serial dependence of market model residuals.
2. **Momentum** is composed of a cumulative 12-month relative strength variable and the historic alpha from the 60-month regression of the security excess return on the S&P 500 excess return.
3. **Size** is the log of the security market capitalization.
4. **Size Nonlinearity** is the cube of the log of the security market capitalization.
5. **Trading Activity** is composed of annualized share turnover of the past 5 years, 12 months, quarter, and month, and the ratio of share turnover to security residual variance.
6. **Growth** is composed of the growth in total assets, 5-year growth in earnings per share, recent earnings growth, dividend payout ratio, change in financial leverage, and analyst-predicted earnings growth.
7. **Earnings Yield** is composed of consensus analyst-predicted earnings to price and the historic earnings to price ratios.
8. **Value** is measured by the book to price ratio.
9. **Earnings Variability** is composed of the coefficient of variation in 5-year earnings, the variability of cash flow, and the variability of analysts' forecasts of earnings to price.
10. **Leverage** is composed of market and book value leverage, and the senior debt ranking.

Long-Term Version, which featured 12 factors, including beta, momentum, size, earnings yield, and growth. An asset manager could prefer to employ a price momentum or earnings growth tilt in the portfolio, and in this case, only the momentum, growth, earnings yield, and size exposures might (should) be non-zero.[19] The main advances of USE4 are:

1. An innovative eigenvector risk adjustment that improves risk forecasts for optimized portfolios by reducing the effects of sampling error on the factor covariance matrix;
2. A Volatility Regime Adjustment designed to calibrate factor volatilities and specific risk forecasts to current market levels;
3. The introduction of a country factor to separate the pure industry effect from the overall market and provide timelier correlation forecasts;
4. A new specific risk model based on daily asset-level specific returns;
5. A Bayesian adjustment technique to reduce specific risk biases due to sampling error;
6. A uniform responsiveness for factor and specific components, providing greater stability in sources of portfolio risk;
7. A set of multiple industry exposures based on GICS®;
8. An independent validation of production code through a double-blind development process to assure consistency and fidelity between research code and production code;
9. A daily update for all components of the model.

The Menchero and Nagy (2015) chapter in this volume employs the USE4 and GEM3 model to study the effectiveness of portfolio construction emphasizing price momentum and earnings yield factor in US and Global equity portfolios.[20] The GEM3 and USE4 risk factors are described in the Appendix.

---

11. **Currency Sensitivity** is composed of the relationship between the excess return on the stock and the excess return on the S&P 500 Index. These regression residual returns are regressed against the contemporaneous and lagged returns on a basket of foreign currencies.
12. **Dividend Yield** is the Barra-predicted dividend yield.
13. **Non-Estimation Universe Indicator** is a dummy variable which is set equal to zero if the company is in the Barra estimation universe and equal to one if the company is outside the Barra estimation universe.

[19]Jose Menchero, D.J. Orr, and Jun Wang, "The Barra US Equity Model (USE4): Methodology Notes, August 2011.

[20]The Barra Global Equity Model, GEM2, offered statistically significant results for optimized Value, Momentum, Liquidity, and Size risk factor portfolios. One needed to have a unit exposure to the particular factor, and zero exposures to all other factors. See Jose Menchero, Andrei Morozov, and John Guerard, "Capturing Equity Risk Premia," in C.F. Lee, J. Finnerty. J. Lee, A.C. Lee, and D. Wort, Security Analysis, Portfolio Management, and Financial Derivatives (Singapore: World Scientific, 2013, Chapter 25).

1.3 The Axioma Risk Model: Fundamental and Statistical Risk Models

The Axioma Robust Risk Model[21] is a multi-factor risk model, in the tradition of the Barra model and Eq. (1.23). Axioma offers both US and World Fundamental and Statistical Risk Models. The Axoma Risk Models use several statistical techniques to efficiently estimate factors. The ordinary least squares residuals (OLS) of Eq. (1.16) are not homoskedastic; that is, when one minimizes the sum of the squared residuals to estimate factors using OLS, one finds that large assets exhibit lower volatility than smaller assets. A constant variance of returns is not found. Axioma uses a weighted least squares (WLS) regression, which scales the asset residual by the square root of the asset market capitalization (to serve as a proxy for the inverse of the residual variance). Robust regression, using the Huber M Estimator, addresses the issue and problem of outliers. (Asymptotic) Principal components analysis (PCA) is used to estimate the statistical risk factors. Ross (1976), Ross and Roll (1980) developed and estimated multi-factor models. Dhrymes, Friend and Gultekin (1984), and Dhrymes, Friend, Gultekin, and Gultekin (1985) estimated that 4–5 factors were presented in size-created portfolios. A subset of assets is used to estimate the factors and the exposures and factor returns are applied to other assets.

Axioma has pioneered two techniques to address the so-called underestimation of realized tracking errors, particularly during the 2008 Financial Crisis. The first technique, known as the Alpha Alignment Factor, AAF, recognizes the possibility of missing systematic risk factors and makes amends to the greatest extent that is possible without a complete recalibration of the risk model that accounts for the latent systematic risk in alpha factors explicitly. In the process of doing so, AAF approach not only improves the accuracy of risk prediction, but also makes up for the lack of efficiency in the optimal portfolios. The second technique, known as the Custom Risk Model, CRM, proposes the creation of a custom risk model by combing the factors used in both the expected-return and risk models, which does not address the factor alignment problem that is due to constraints. Several practitioners have decided to perform a "post-mortem" analysis of mean-variance portfolios, attempted to understand the reasons for the deviation of ex-post performances from ex-ante targets, and used their analysis to suggest enhancements to mean-variance optimization inputs, in order to overcome the discrepancy. Lee and Stefek (2008) and Saxena and Stubbs (2012) define this as a factor alignment problem (FAP), which arises as a result of the complex interactions between the factors used for forecasting expected returns, risks and constraints.[22] While predicting expected returns is exclusively a forward-looking activity, risk prediction focuses on

[21] Axioma Robust Risk Model Handbook, January 2010.

[22] The author expresses great appreciation for many conversations with Anureet Saxena on this topic.

explaining the cross-sectional variability of returns, mostly by using historical data. Expected-return modelers are interested in the first moment of the equity return process, while risk modelers focus on the second moments. These differences in ultimate goals inevitably introduce different factors for expected returns and risks. Even for the "same" factors, expected-return and risk modelers may choose different definitions for good reasons. Constraints play an important role in determining the composition of the optimal portfolio. Most real-life quantitative strategies have other constraints that model desirable characteristic of the optimal portfolio. For example, a client may be reluctant to invest in stocks that benefit from alcohol, tobacco or gambling activities on ethical grounds, or may constrain their portfolio turnover so as to reduce their tax burden.

The naïve application of the portfolio optimization has the unintended effect of magnifying the sources of misalignment. The optimized portfolio underestimates the unknown systematic risk of the portion of the expected returns that is not aligned with the risk model. Consequently, it overloads the portion of the expected return that is uncorrelated with the risk factors. The empirical results in a test-bed of real-life active portfolios based on client data show clearly that the above-mentioned unknown systematic risk is a significant portion of the overall systematic risk, and should be addressed accordingly. Saxena and Stubs (2012) reported that the earning-to-price (E/P) and book-to-price (B/P) ratios used in USER Model and Axioma Risk Model have average misalignment coefficients of 72 % and 68 %, respectively. While expected-return and risk models are indispensable components of any active strategy, there is also a third component, namely the set of constraints that is used to build a portfolio. Saxena and Stubbs (2012) proposed that the risk variance-covariance matrix C be augmented with additional auxiliary factors in order to complete the risk model. The augmented risk model has the form of

$$C_{\text{new}} = C + \sigma_{\underline{\alpha}}^2 \underline{\alpha} \cdot \underline{\alpha}' + \sigma_{\underline{\gamma}}^2 \underline{\gamma} \cdot \underline{\gamma}', \tag{1.28}$$

where $\underline{\alpha}$ is the alpha alignment factor (AAF), σ_α is the estimated systematic risk of $\underline{\alpha}$, $\underline{\gamma}$ is the auxiliary factor for constrains, and σ_γ is the estimated systematic risk of $\underline{\gamma}$. The alpha alignment factor $\underline{\alpha}$ is the unitized portion of the uncorrelated expected-return model, i.e., the orthogonal component, with risk model factors. Saxena and Stubbs (2012) reported that the AAF process pushed out the traditional risk model-estimated efficient frontier. Saxena and Stubbs (2015) refer to as alpha in the augmented regression model as the implied alpha. According to Saxena and Stubbs (2015), the base risk model, BRM, assumes that any factor portfolio uncorrelated with X-common risk factors has only idiosyncratic risk. Z is the exposure matrix associated with systematic risk factors missing from the base risk model, and the risk model fails to account for the systematic risk of portfolios with exposure to the Z factors. Saxena and Stubbs (2015) report that there is a small increment to specific risk compared to its true systematic risk.

Saxena and Stubbs (2012) applied their AAF methodology to the USER model, running a monthly backtest based on the above strategy over the time period 2001–2009 for various tracking error values of σ chosen from $\{4\,\%, 5\,\% \ldots\ 8\,\%\}$. For each value of σ, the backtests were run on two setups, which were identical in all respects except one, namely that only the second setup used the AAF methodology ($\sigma_\alpha = 20\%$). Axioma's fundamental medium-horizon risk model (US2AxiomaMH) is used to model the active risk constraints. Saxena and Stubbs (2012) analyzed the time series of misalignment coefficients of alpha, implied alpha and the optimal portfolio, and found that almost 40–60 % of the alpha is not aligned with the risk factors. The alignment characteristics of the implied alpha are much better than those of the alpha. Among other things, this implies that the constraints of the above strategy, especially the long-only constraints, play a proactive role in containing the misalignment issue. In addition, not only do the orthogonal components of both the alpha and the implied alpha have systematic risk, but the magnitude of the systematic risk is comparable to that of the systematic risk associated with a median risk factor in US2AxiomMH. Saxena and Stubbs (2012) showed the predicted and realized active risks for various risk target levels, and noted the significant downward bias in risk prediction when the AAF methodology is not employed.[23] The realized risk–return frontier demonstrates that not only does using the AAF methodology improve the accuracy of the risk prediction, it also moves the ex-post frontier upwards, thereby giving ex-post performance improvements. In other words, the AAF approach recognizes the possibility of missing systematic risk factors and makes amends to the greatest extent that is possible without a complete recalibration of the risk model that accounts for the latent systematic risk in alpha factors explicitly. In the process of doing so, AAF approach not only improves the accuracy of risk prediction, but also makes up for the lack of efficiency in the optimal portfolios.[24] Saxena and Stubbs (2015) extended their 2012 Journal of Investing research and reported positive frontier spreads.

As a further test of the applicability of the Axioma statistical and fundamental risk models in analyzing US stocks, let us analyze a set of all publically listed US

[23]The bias statistic shown is a statistical metric that is used to measure the accuracy of risk prediction; if the ex-ante risk prediction is unbiased, then the bias statistic should be close to 1.0. Clearly, the bias statistics obtained without the aid of the AAF methodology are significantly above the 95 % confidence interval, which shows that the downward bias in the risk prediction of optimized portfolios is statistically significant. The AAF methodology recognizes the possibility of inadequate systematic risk estimation, and guides the optimizer to avoid taking excessive unintended bets.

[24]Guerard, Markowitz, and Xu (2013, 2015) created efficient frontiers using both of the AxiomaRisk Models, and found that the statistically based Axioma Risk Model, the authors denoted as "STAT," produced higher geometric means, Sharpe ratios, and information ratios than the Axioma fundamental Risk Model, denoted as "FUND." The AAF technique was particularly useful with composite models of stock selection using fundamental data, momentum, and earnings expectations data. Furthermore, the geometric means and Sharpe ratios increase with the targeted tracking errors; however, the information ratios are higher in the lower tracking error range of 3–6 %, with at least 200 stocks, on average, in the optimal portfolios. The Guerard et al. studies assumed 150 basis points, each way, of transactions costs. The use of ITG cost curves produced about 115–125 basis points of transactions costs, well under the assumed costs. The Guerard et al. studies also used the Sungard APT statistical model which produced statistical significant asset selection in US and global portfolios.

stocks for the 1999–2013 time period. We test two models, the earnings forecasting strategy variable, CTEF, of Guerard, Gultekin, and Xu (1997) and Guerard et al. (2015), and a ten-factor regression-based US expected returns strategy, USER, tested in Guerard, Gultekin, and Xu (2012) and Saxena and Stubbs (2012). We simulate US stocks covered by I/B/E/S with the same stipulated conditions: (1) a 4 % maximum position; (2) 8 % monthly turnover; a 35 basis point threshold position; and 150 basis points of transactions costs. The two strategies produce similar results with the respective Axioma risk models and techniques: (1) both CTEF and USER produce higher active returns as targeted racing errors rise; (2) active returns and Information Ratios, the portfolio excess return divided by the portfolio tracking error, TE, are higher with the Axioma statistical model relative to the Axioma fundamental model; and (3) the Alpha Alignment Factor (AAF) of 20 % portfolios produce raises CTEF IRs with both Axioma risk models relative to non-AAF portfolios with targeted tracking errors of 5, 8, and 9 % (see Table 1.1).

Table 1.1 Investment variables and Axioma risk models

Axioma risk models						
WRDS US Backtest Universe, 1999–2013						
		CTEF			USER	
TE	Active risk	Active returns	IRs	Active risk	Active returns	IRs
Axioma world fundamental risk model						
4	4.36	3.56	0.816	4.16	2.54	0.551
5	5.41	3.57	0.660	5.82	2.74	0.470
6	6.28	3.43	0.546	6.88	3.17	0.462
7	6.82	3.94	0.578	7.81	2.13	0.273
8	7.72	5.03	0.651	8.56	2.37	0.277
9	8.05	4.55	0.565	9.53	3.08	0.323
10	8.48	4.82	0.569	10.08	3.73	0.370
Axioma world statistical risk model						
4	4.70	4.73	1.007	4.71	3.56	0.754
5	5.87	4.57	0.778	6.05	3.99	0.660
6	6.57	5.26	0.790	7.14	3.86	0.540
7	7.66	5.03	0.703	8.00	3.93	0.491
8	8.00	6.53	0.816	8.78	3.94	0.449
9	8.43	5.89	0.698	9.44	4.08	0.432
10	9.12	6.38	0.700	10.28	4.58	0.446
Axioma world statistical risk model, AAF 20 %						
4	3.61	2.22	0.642	3.41	3.04	0.893
5	4.40	3.76	0.854	4.58	3.58	0.780
6	5.29	3.92	0.753	5.85	3.06	0.524
7	6.32	4.20	0.665	7.10	3.38	0.476
8	6.83	6.41	0.938	7.45	3.83	0.514
9	7.67	5.59	0.729	8.70	3.74	0.429
10	8.35	5.57	0.667	9.08	4.94	0.544

One sees in Table 1.1 that the CTEF variable produces higher active returns than the ten-factor composite model. A similar result was reported for global stocks in Guerard, Markowitz, and Xu (2015).

Let us extend the previous analysis of CTEF to US stocks for 1999–September 2014. We simulate US stocks covered by I/B/E/S in the USER database, 1999-September 2014 with the same simulation conditions: (1) a 4 % maximum position; (2) 8 % monthly turnover; a 35 basis point threshold position; and 150 basis points of transactions costs. We report simulation results in Table 1.2.

Many of the conclusions a reported for global stock reported in Guerard et al. (2015): (1) Geometric Means and Information Ratios favor the Statistical Risk Models relative to the Fundamental Risk Models; (2) Geometric Means and Information Ratios favor the Statistical Risk Models relative to the Fundamental Risk Models with an Alpha alignment Factor of 20 %; (3) a names constraint, of 70 stocks, actually enhances portfolio Information Ratios; and (4) one can use only one factor, F1, and still generate positive and economically meaningful returns. The reader is referred to Chart 1.1 for a plot of the Markowitz Efficient Frontiers of CTEF using the Axioma Statistical and Fundamental Risk models, with and without AAF. The use of the Alpha Alignment Factor of 20 %, consistent with Saxena and Stubbs (2012), pushed out the frontiers.[25] Yes, Virginia, a one-factor model Treynor Model works, particularly at lower, 4–6 % targeted tracking errors. A four factor model, using the first four factors of the Axioma Statistical Risk Model, reflecting the work of Phoebus Dhrymes, Irwin Friend, and Gultekin (1984) and a later paper where they were joined by and Mustafa Gultekin (1985) works well in the 1999–September 2014 period, but not as well as the 15 factors of the Axioma Statistical Risk Model, using the names constraints, see Chart 1.2. Thus, one needs the Axioma PCA-based factors.

Sivaramakrishnan and Stubbs (2013) proposed the creation of an Axioma custom risk model by combing the factors used in both the expected-return and risk models, which does not address the factor alignment problem that is due to constraints. The Sivaramakrishnan and Stubbs model allowed great interaction with clients to produce several variations on risk models that were consistent with particular clients' needs of risk exposures. Ceria, Sivaramakrishnan and Stubbs (2015) contributed a chapter to this volume which illustrates how each alpha signal can be transformed into a factor mimicking portfolio, and how the alpha signals can be combined into a target portfolio with a mean-variance optimization (MVO) problem. The new (combined) alpha signal is constructed as the implied alpha of the target portfolio and used with a custom risk model. The portfolio is consistent as it satisfies the relevant implementation constraints.

[25] As the author was revising this chapter, he was delighted to see the Saxena and Stubbs (2015) AAF analysis.

Table 1.2 Axioma fundamental and statistical risk models of forecasted earnings acceleration, CTEF, in the US Expected Returns (USER) stocks database, 1999–2014

Risk models	Targeted tracking error	Annualized returns	Annualized standard deviation	Annualized active returns	Annualized active risk	Sharpe ratio	Information ratio	Number of stocks	AAF	Number of selected factors
FUND	4	10.25	19.24	3.56	4.36	0.537	0.816	1469		
	5	10.27	19.53	3.57	5.41	0.526	0.660	1467		
	6	10.13	19.97	3.43	6.28	0.507	0.541	1518		
	7	10.63	20.15	3.14	6.82	0.528	0.578	1523		
	8	11.72	20.90	5.03	7.72	0.567	0.651	1551		
	9	11.24	20.90	4.55	8.05	0.538	0.565	1561		
	10	11.52	20.65	4.82	8.48	0.558	0.569	1569		
STAT	4	10.70	19.46	3.99	4.80	0.393	0.832	1522		
	5	11.24	19.96	4.53	5.80	0.410	0.781	1531		
	6	11.89	20.30	5.19	6.60	0.435	0.786	1549		
	7	12.01	20.48	5.30	7.19	0.437	0.737	1547		
	8	13.34	21.06	6.63	8.11	0.488	0.817	1566		
	9	12.48	20.54	5.77	8.64	0.454	0.667	1582		
	10	12.99	20.55	6.28	9.00	0.483	0.698	1586		

(continued)

Table 1.2 (continued)

Risk models	Targeted tracking error	Annualized returns	Annualized standard deviation	Annualized active returns	Annualized active risk	Sharpe ratio	Information ratio	Number of stocks	AAF	Number of selected factors
FUND	4	12.07	19.47	5.36	7.89	0.463	0.680	1602	20	
	5	12.83	19.73	6.12	8.41	0.495	0.728	1598		
	6	13.61	20.02	6.90	9.23	0.527	0.748	1604		
	7	14.24	20.27	7.53	9.55	0.552	0.789	1597		
	8	14.24	20.42	7.53	9.57	0.547	0.787	1588		
	9	DNF								
	10	14.04	20.67	7.33	9.76	0.531	0.751	1600		
STAT	4	12.39	19.34	5.60	7.75	0.483	0.724	1586	20	
	5	13.09	20.01	6.38	8.71	0.501	0.733	1603		
	6	13.39	19.87	6.62	8.99	0.517	0.736	1602		
	7	13.53	20.24	6.82	9.50	0.518	0.718	1598		
	8	13.78	20.49	7.08	9.55	0.530	0.741	1582		
	9	13.32	20.23	6.61	9.67	0.507	0.684	1591		
	10	13.29	20.50	6.50	9.60	0.489	0.680	1594		
STAT	4	9.64	18.68	2.93	3.41	0.352	0.859	70	20	
	5	10.37	19.02	3.66	4.75	0.385	0.772	70		
	6	10.72	19.38	4.01	5.72	0.396	0.702	70		
	7	12.80	19.54	6.09	6.37	0.498	0.956	70		
	8	11.51	19.95	4.80	6.96	0.242	0.689	70		
	9	12.51	19.70	5.80	7.48	0.480	0.781	70		
	10	13.44	20.38	6.73	8.20	0.510	0.822	70		

STAT	4	10.40	18.40	3.69	4.74	0.399	0.778	70	20	F1
	5	11.63	18.77	4.92	5.42	0.457	0.909	70		
	6	12.93	18.85	6.22	6.85	0.524	0.908	70		
	7	11.51	19.57	4.80	8.18	0.432	0.587	70		
	8	13.44	19.40	6.73	8.84	0.535	0.762	70		
	9	12.89	19.42	6.18	9.94	0.506	0.622	70		
	10	11.85	19.49	5.14	10.18	0.451	0.505	70		
STAT	4	9.29	18.84	2.58	4.10	0.331	0.629	70	20	F1, F2, F3, F4
	5	11.11	18.84	4.40	5.50	0.427	0.799	70		
	6	10..94	18.83	4.23	6.42	0.419	0.660	70		
	7	13.80	19.72	7.10	7.58	0.545	0.936	70		
	8	13.13	19.60	6.42	8.46	0.514	0.764	70		
	9	12.46	20.23	5.76	9.08	0.465	0.634	70		
	10	13.46	20.69	6.75	9.70	0.501	0.693	70		

Alpha Alignment Factor (AAF) analysis of 20 % with a names constraint

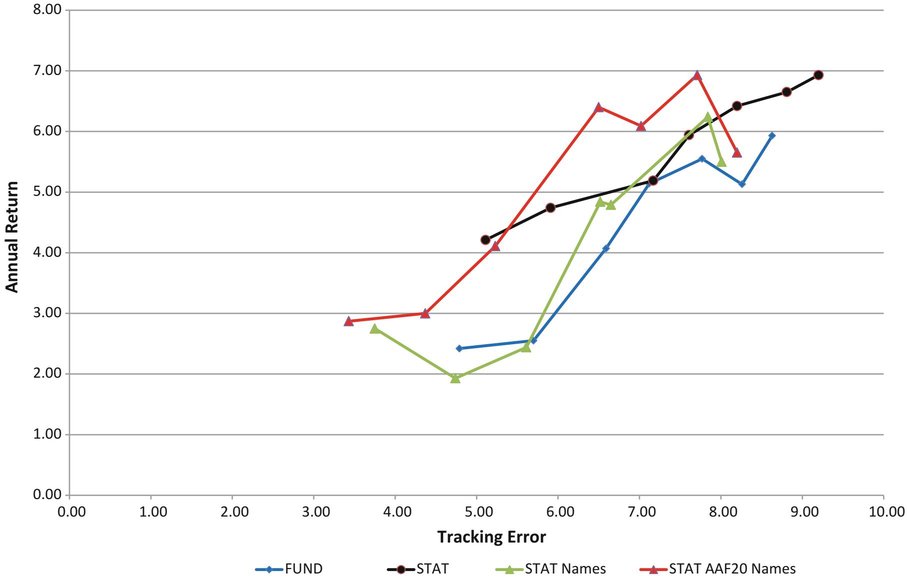


Chart 1.1 CTEF return and tracking error trade-off comparing Axioma risk models

Chart 1.2 CTEF return and tracking error analysis using the Axioma World vs US risk models

1.4 Assessing Mutual Funds: The Treynor Index and Other Measurement Techniques

This section draws heavily from Lee et al. (2010) and reviews the selectivity, market timing, and overall performance of equity funds in the USA during January 1990 until September 2005.[26] Lee, Lee, and Liu (denoted "LLL") used Sharpe, Treynor, and Jensen measures to evaluate the selectivity performance of mutual fund managers. In addition, we also used the Treynor–Mazuy and Lee–Rahman models to evaluate the selectivity and timing performance of mutual fund managers. Based upon these measures and models, LLL reported that about one-third of funds had significantly positive selectivity ability, and some had timing ability for the mutual fund managers. Nevertheless, without considering transaction costs and taxes, the actual investment for most mutual funds compared to a passive investment strategy still appears to take the lead.

The investment of mutual funds has been extensively studied in finance. Over the last few decades, there has been a dramatic increase in the development of instruments measuring the performance of mutual funds. Early researchers (Treynor (1965), Sharpe (1966), and Jensen (1968)) employed a one parameter indicator to evaluate the portfolio performance. However, these studies assumed the risk levels of the examined portfolios to be stationary through time. Fama (1972) and Jensen (1972) pointed out that the portfolio managers may adjust their risk composition according to their anticipation for the market. Moreover, Fama (1972) suggested that the managers' forecasting skills can be divided into two parts: the selectivity ability and the market timing ability. The former is also named as micro-forecasting, involving the identification of the stocks that are undervalued or overvalued relative to the general stocks. The latter is also named as macro-forecasting, involving the forecast of future market return. In other words, the selectivity and market timing abilities of fund managers are viewed as important factors deciding the overall fund performance.[27]

Treynor and Mazuy (1966) used a quadratic term of the excess market return to test for market timing ability. It can be viewed as the extension of the Capital Asset Pricing model (CAPM). If the fund manager can forecast market trend, he will change the proportion of the market portfolio in advance. Jensen (1972) developed the theoretical structure for the timing ability. Under the assumption of a joint normal distribution of the forecasted and realized returns, Jensen showed that the correlation between the managers' forecast and the realized return can be used to

---

[26]C. F. Lee, A. Lee, and N. Liu, "Alternative Model to Evaluate Selectivity and Timing Performance of Mutual Fund Managers: Theory and Evidence," in J. Guerard, Jr. Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques (New York: Springer, 2010).

[27]However, Brinson, Singer, and Beebower (1991) found that selectivity and market timing abilities only have small influence on fund performance (<10 %). The overall performance should be mostly decided by asset allocation between stock and bond markets.

measure the timing ability. Bhattacharya and Pfleiderer (1983) extended Jensen's (1972) work and used a simple regression method to obtain accurate measures of selectivity and market timing ability. Lee and Rahman (1990) further corrected the inefficient estimated of parameters by a Generalized Least Squares (GLS) method. In addition, Henriksson and Merton (1981) used options theory, developed by Merton to explain the timing ability.

Lee et al. (2010) empirically examined the mutual fund performance by using six models, proposed respectively by Treynor (1965), Sharpe (1966), Jensen (1968), Treynor and Mazuy (1966), Henriksson and Merton (1981), and Lee and Rahman (1990). The monthly returns for 189 months (January 1990 to September 2005) for a sample of 628 open-end equity funds were used. Let us introduce the reader to the basic ratios and parameter used in evaluating portfolio performance.

Treynor (1965) uses the concept of the security market line[28] drawn from the CAPM to get a coefficient β. Under the assumption of complete diversification of asset allocation, it means that we just have systematic risk measured by β. The Treynor index (TI) measuring the reward per unit of systematic risk for the portfolio can be showed as follows:

$$\mathrm{TI} = \frac{\bar{r}_\mathrm{p} - r_f}{\beta_\mathrm{p}}, \tag{1.28}$$

where $\bar{r}_\mathrm{p}$ is the average return of the pth mutual fund, and r_f is defined as risk-free rate. The numerator of Treynor index can be viewed as excess return on the portfolio. This ratio is a risk-adjusted performance value. This indicator is suitable for valuing the performance of a well-diversified portfolio; this is because it just takes the systematic risk into account.

Different from Treynor (1965) and Sharpe (1966) argues the phenomenon that the fund managers will be in favor of fewer stocks. Therefore, it is impossible to diversify the individual risks completely. In other words, the excess return should be calculated based on the total risk (including systematic and nonsystematic risks). The Sharpe Index (SI), applying the concept of the capital market line[29] can be written as:

$$\mathrm{SI} = \frac{\bar{r}_\mathrm{p} - r_f}{\sigma_\mathrm{p}}, \tag{1.29}$$

where σ_p is the standard deviation of the portfolio, namely total risk. The Sharpe index is expressed as the reward per unit of total risk. The higher the two indices mentioned above, the better the fund's performance. Because this measure is based on the total risk, it enables to measure the performance of the portfolio which is not very diversified.

[28] At equilibrium, all assets are located on this line.

[29] In the presence of a risky asset, this straight line is the efficient frontier for all investors.

Jensen (1968) proposes a regression-based view to measure the performance of the portfolio. The Jensen index (or called Jensen alpha) utilizes the CAPM to determine whether a fund manager outperformed the market. It's formula is as follows:

$$R_{\mathrm{p},t} = \alpha_{\mathrm{p}} + \beta_{\mathrm{p},t} R_{m,t} + u_{\mathrm{p},t}, \tag{1.30}$$

where $R_{\mathrm{p},t}$ and $R_{m,t}$ are the excess returns $\left(R_t = r_t - r_f\right)$ at time t of the portfolio return and the market return, respectively. The term of $u_{\mathrm{p},t}$ in the formula is the residual at time t. The coefficient α_{p} is used to measure the performance of mutual funds in the sense of the additional return due to the manager's choice. It also represents the fund manager's selectivity ability without considering timing ability. A significantly positive and high value of Jensen alpha indicates superior performance compared with the market index.

Treynor and Mazuy (1966), putting a quadratic term of the excess market return into Eq. (1.28), provide us with a better framework for the adjustments of the portfolio's beta to test a fund manager's timing ability. The fund manager with timing ability will be able to adjust the risk exposure from the market. To take a simple example, if a fund manager expects a coming up (down) market, he will hold a larger (smaller) proportion of the market portfolio. Therefore, the portfolio return can be viewed as a convex function of the market return. The equation can be given below:

$$R_{\mathrm{p},t} = \alpha_{\mathrm{p}} + \beta_1 R_{m,t} + \beta_2 R_{m,t}^2 + \varepsilon_{\mathrm{p},t}, \tag{1.31}$$

where the coefficient β_2 is used to measure the timing ability. When β_2 is significantly larger than zero, it represents that, in a up (down) market, the increasing (decreasing) proportion in the risk premium of the mutual fund is larger than that in the market portfolio. This model was formulated empirically by Treynor and Mazuy (1966). It was then theoretically validated by Jensen (1972), and Bhattacharya and Pfleiderer (1983).

Henriksson and Merton (1981) used options theory to explain the timing ability. It consists of a modified version of the CAPM which takes the manager's two objectives into account, and depends on whether he forecasts that the market return will or will not be better than the risk-free asset return. They view the coefficient β as a binary variable. This means that a fund manager with market timing ability should have different β values in the up and down markets.[30] We can express the equation as:

---

[30]Ferson and Schadt (1996) assumed that market prices of securities reflected public information and allowed betas of stocks and portfolios to change with economic conditions. See Christopherson, Carino, and Ferson (2009), Chap. 12 for a more complete discussion of conditional performance evaluation.

$$R_{p,t} = \alpha_p + \beta_1 R_{m,t} + \beta_2 \text{Max}\,(-, R_{m,t}) + \varepsilon_{p,t}. \tag{1.32}$$

If $\beta_2 > 0$, this shows that the manager has the ability to forecast a market to be down or up. For an up market (a down market), the Eq. (1.32) can be expressed as $R_{p,t} = \alpha_p + \beta_1 R_{m,t} + \varepsilon_{p,t}\left(R_{p,t} = \alpha_p + (\beta_1 - \beta_2)\, R_{m,t} + \varepsilon_{p,t}\right)$.

Jensen (1972) showed that the timing ability can be measured by the correlation between the managers' forecast and the realized return. Bhattacharya and Pfleiderer (1983) modify Jensen's (1972) model[31] to propose a regression-based model to evaluate the market timing and selectivity abilities.[32]

Lee et al. (2010) used the alternative methods to examine the selectivity, market timing, and overall performance for the open-end equity mutual funds.[33] The samples used were the monthly returns of the 628 mutual funds[34] ranging from January, 1990 to September, 2005, 189 monthly observations. The fund data were obtained from the CRSP Survivor-Bias-Free US Mutual Fund Database. Then, they used ICDI's fund objective codes to sort the objectives of the mutual funds. In total, there are 23 types of mutual fund objectives; simplifying into as two groups, growth funds and non-growth funds,[35] consisting of 439 growth funds and 189 non-growth funds. In addition to the CRSP fund data, the S&P 500 stock index obtained from Datastream is used for the return of the market portfolio. Moreover, they use the Treasure bill rate with a 3-months holding period as the risk-free return. The Treasury bill rate is available from the website of the Federal Reserve Board.

We present a summary table, Table 1.3, of the Lee et al. (2010) estimates of mutual fund performance.

The difference of monthly returns of growth and non-growth funds is quite substantial. The result for the 1990–2005.09 period shows that 82 % (69 %) of the growth (non-growth) funds have better performance than the market. This seems to point out the growth funds are more valuable to be invested than the non-growth

[31] In a framework of Jensen (1972), the coefficients in the model can't be estimated efficiently. However, with some assumptions proposed by Bhattacharya and Pfleiderer (1983), we can get the efficient estimators. The detail can be found in Lee and Rahman (1990), p. 265–266.

[32] Lee and Rahman (1990) find that the residual terms of the Bhattacharya and Pfleiderer exists heteroscedasticity. Thus, the estimated coefficients are not efficient. The way to solve this problem is to calculate the variance of the residuals, ω_t and τ_t.

Lee and Rahman utilized a GLS method with correction for heteroscedasticity to adjust the weights of the variables in Eqs. (1.8) and (1.10) by σ_ω^2 and σ_t^2.

[33] In addition to equity funds (the code in CRSP is EQ), Standard & Poor's Main Category provides the other four kinds of funds to define fund styles, i.e., fixed income (FI), money market (DG), asset allocation (AA), and convertible (CT), but they are not analyzed in this study.

[34] We delete the funds with any missing values in this period. In addition, the funds with over 20 zero returns are also dropped. This is because we view too many zero values as missing data or lower liquidity for the fund. The list of the fund names is available from the authors on request.

[35] Three types of mutual funds belongs to the growth group, i.e., aggressive growth (AG), growth and income (GI), and long term growth (LG); the other 20 types are put in the non-growth group.

Table 1.3 Summary table of mutual fund performance measurement, adapted from Lee et al. (2010)

628 mutual funds, 1990–2005.09			
Measurement	Growth funds	Non-growth funds	Index
Treynor Index	0.0052	0.0045	0.0032
Sharpe Index	0.1037	0.0824	0.0774
Jensen Index	0.0017	0.0013	
Treynor–Mazuy	0.0035	0.0042	
Henriksson–Merton			
Selectivity	0.0049	0.0069	
Timing	−0.1956	−0.3435	
Lee–Rahman			
Selectivity	0.0033	0.0041	
Timing	0.0816	0.0982	

funds.[36] Kosowski, Timmermann, Wermers, and White (2006) reported that the top 5 % of abnormal-growth funds produced persistent, and statistically significant excess returns, relative to the four-factor Fama-French (1992) and Carhart (1997) risk factors, for 2118 open-end US equity funds for the January 1975–December 2002 period. Persistent was determined by two periods, 1975–1989 and 1990–2002 periods. Moreover, Kosowski et al. (2006) reported that the worst two decile of abnormal-growth significantly underperformed the universe of funds. They found no evidence of "stars" or significant and persistent outperformance in income funds.

The performance of the Treynor index is mainly based on the systematic risk obtained from the CAPM. It is appropriate to evaluate the well-diversified portfolio. Different from the Treynor index, the Sharpe index takes account of the total risk of the portfolio. The performance measured by the Sharpe index is also given in Table 1.3. Generally speaking, the Sharpe Index results are consistent with the Treynor Index results in Table 1.3. Unlike the first two measures, the Jensen index is calculated by carrying out a simple regression. The Jensen alpha is also based on the CAPM and measures the share of additional return that is due to the manager's choices. Most of the funds have positive Jensen alphas. Lee et al. (2010) reported that 30 % of the growth funds for the entire period are significantly larger than zero with 95 % confidence.

The first three indicators assume that the portfolio risk is stationary and only take the stock selection into account. If we want to modify the level of the portfolio's exposure to the market risk, the timing ability has to be adopted. The models with the ability to test market timing are against the CAPM. This is due to permit variations in the portfolio's beta over the investment period. There are three models for evaluating the selectivity and timing abilities in this study. The Treynor–Mazuy

[36] LLL did not consider the transaction costs and taxes here. In general, the growth fund will ask for a higher commission than the non-growth fund.

model, a quadratic version of the CAPM, tests market timing abilities. Lee, Lee, and Liu reported that over 80 % of them have negative values of timing ability and a considerable ratio among them have significantly negative estimates. For the selectivity ability, a very high ratio of the funds has positive estimates and many of them are significantly positive. Only very few funds have significantly negative estimates. The Henriksson–Merton model is also a modified version of the CAPM which applies a binary choice for the manager. It depends on the market return will or will not perform better than the risk-free asset return. The estimates of the selectivity ability even display larger values. Furthermore, the correlations between the estimates of timing and selectivity ability are −0.85 for the entire period. The Lee–Rahman model assumes no negative timing ability. The non-growth funds have better timing ability than the growth funds for all periods. Different from the previous two models, the correlations between the estimates of timing and selectivity ability are 0.43 for the entire period. Moreover, 40 of the funds in the entire period have significantly positive estimates in both selectivity and timing ability.

What do we know about mutual funds and their relative performance? Bogle (2000, 2009) tells us that actively managed mutual funds underperform the market by approximately 200 basis points. Lee et al. (2010) reported that growth funds outperformed the market for the 1995–2005 period. Kosowski et al. (2006) and Wei, Wermers, and Yao (2016) reported that contrarian funds outperformed the market during the 1995–2012 period. Berk and van Binsbergen (2016a, 2016b) demonstrate that the Treynor CAPM model dominates the Fama-French (1992) and Carhart (1997) models in assessing mutual fund performance for the 1995–2011 period. Thus, after 50 years, the work of Jack Treynor and his ranking of mutual funds is still not only relevant, but dominant, in many studies.

1.5 Conclusions and Summary

This chapter addresses several aspects of risk, return, and performance measurement. It is important to see how quantitative analysis was developed in investment research and analysis. Harry Markowitz, Bill Sharpe, Jan Mossin, and Jack Treynor pioneered capital market equilibrium and the creation and estimation of the Capital Asset Pricing Model. In the 1970s, Barr Rosenberg and his colleagues at Barra developed and estimated multi-factor risk models. Recent research and commercialization by Axioma has furthered portfolio optimization. Mutual fund performance measurement, pioneered by Sharpe and Treynor, incorporates risk estimation.

USE4 Descriptor Definitions

The definitions of the descriptors which underlie the risk indices in USE4 (L, denoting long-term version). The method of combining these descriptors into risk indices is proprietary to Barra. The USE4 model starts in June 1995 and uses the US component of MSCI All Country World Investable Market Index (IMI).

1. **Beta** is the historic beta;
2. **Non-Linear Beta** is the non-linear historic beta;
3. **Residual Volatility** is composed of variables explaining returns of high-volatility stocks not associated with the estimated beta factor including the daily standard deviation;
4. **Momentum** is composed of a cumulative 6–12-month relative strength variable;
5. **Size** is the log of the security market capitalization;
6. **Non-Linear Size** is the cube of the log of the security market capitalization;
7. **Liquidity** is composed of annualized share turnover of the past 12 months, quarter, and month;
8. **Growth** is composed of the growth in total assets, 5-year growth in earnings per share, sales growth, and analyst-predicted sales growth;
9. **Earnings Yield** is composed of consensus analyst-predicted earnings to price and the historic earnings to price ratios, and cash earnings-to-price ratio;
10. **Book-to-Price** is measured by the book to price ratio;
11. **Leverage** is composed of the debt-to-assets ratio, and market and book value leverage;
12. **Dividend Yield** is the dividend-to-price ratio.

The GEM3 risk factor indices are identical to the USE4 factor risk indices with the omission of Non-Linear Beta.

References

Berk, J. B., & van Binsbergen, J. H. (2016a). Assessing asset pricing models using revealed preference. *Journal of Financial Economics, 119*, 1–23.

Berk, J. B., & van Binsbergen, J. H. (2016b). Leveling the playing field. In *Portfolio construction, measurement and efficiency*. New York: Springer.

Bhattacharya, S., & Pfleiderer, P. (1983). *A note on performance evaluation* (Technical Report 714). Stanford, CA: Stanford University, Graduate School of Business.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. Jensen (Ed.), *Studies in the theory of capital markets*. New York: Praeger.

Bogle, J. C. (2000). *Common sense on mutual funds: New imperatives for the intelligent investor*. New York: John Wiley & Sons.

Bogle, J. C. (2009). *Common sense on mutual funds*. New York: John Wiley & Sons.

Brinson, G. P., Singer, B. D., & Beebower, G. L. (1991). Determinants of portfolio performance II: An update. *Financial Analysts Journal, 47*, 40–48.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.

Ceria, S., Sivaramakrishnan, K., & Stubbs, R. A. (2015). Alpha construction in a consistent investment process. In *Portfolio construction, measurement and efficiency*. New York: Springer.

Christopherson, J. A., Carino, D. R., & Ferson, W. E. (2009). *Portfolio performance measurement and benchmarking*. New York: McGraw-Hill Companies.

Dhrymes, P. J., Friend, I., & Gultekin, N. B. (1984). A critical re-examination of the empirical evidence on the arbitrage pricing theory. *Journal of Finance, 39*, 323–346.

Dhrymes, P. J., Friend, I., Gültekin, M. N., & Gültekin, N. B. (1985). New tests of the APT and their implications. *Journal of Finance, 40*, 659–674.

Elton, E. J., & Gruber, M. J. (1970). Homogeneous groups and the testing of economic hypothesis. *Journal of Financial and Quantitative Analysis, 5*, 581–602.

Fama, E. F. (1972). Components of investment performance. *Journal of Finance, 27*, 551–567.

Fama, E. F., & French, K. R. (1992). Cross-sectional variation in expected stock returns. *Journal of Finance, 47*(2), 427–465.

Fama, E. F., & French, K. R. (1995). Size and the book-to-market factors in earnings and returns. *Journal of Finance, 50*(1), 131–155.

Farrell, J. L., Jr. (1974). Analyzing covariance of returns to determine homogeneous stock groupings. *Journal of Business, 47*, 186–207.

Farrell, J. L., Jr. (1997). *Portfolio management: Theory and applications*. New York: McGraw-Hill/Irwin.

Ferson, W., & Schadt, R. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance, 51*, 425–462.

Grinold, R., & Kahn, R. (2000). *Active portfolio management*. New York: McGraw-Hill/Irwin.

Grinold, R., Rudd, A., & Stefek, D. (1989). Global factor: Factor friction? *Journal of Portfolio Management*.

Guerard, J. B. (2010). *The handbook of portfolio construction: Contemporary applications of Markowitz techniques*. New York: Springer. Chapter 3.

Guerard, J. B., Jr., Gultekin, M., & Stone, B. K. (1997). The role of fundamental data and analysts' earnings breadth, forecasts, and revisions in the creation of efficient portfolios. In A. Chen (Ed.), *Research in finance* vol. 15.

Guerard, J. B. (2010). *The handbook of portfolio construction: Contemporary applications of Markowitz techniques*. New York: Springer. Chapter 3.

Guerard, J. B., Jr., Markowitz, H. M., & Xu, G. (2013). Global stock selection modeling and efficient portfolio construction and management. *Journal of Investing, 22*, 121–128.

Guerard, J. B., Jr., Markowitz, H. M., & Xu, G. (2015). Earnings forecasting in a global stock selection model and efficient portfolio construction and management. *International Journal of Forecasting, 31*, 550–560.

Guerard, J. B., Jr., Rachev, S. T., & Shao, B. P. (2013). Efficient global portfolios: Big data and investment universes. *IBM Journal of Research and Development, 57*, 11.1–11.11.

Guerard, J. B., Jr., Xu, G., & Gultekin, M. N. (2012). Investing with momentum: The past, present, and future. *Journal of Investing, 21*, 68–80.

Gültekin, M. N., & Gültekin, N. B. (1987). Stock return anomalies and tests of APT. *Journal of Finance, 42*, 1213–1224.

Henriksson, R. D., & Merton, R. C. (1981). On market timing and investment performance. . Statistical procedure for evaluating forecasting skills. *Journal of Business, 54*, 513–534.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *Journal of Finance, 23*, 389–416.

Jensen, M. C. (1972). Optimal utilization of market forecasts and the evaluation of investment performance. In G. P. Szego & S. Karl (Eds.), *Mathematical methods in investment and finance*. Amsterdam: Elsevier.

Kosowski, R., Timmermann, A., Wermers, R., & White, H. (2006). Can mutual fund "stars" realypick stocks? New evidence from a bootstrap analysis. *Journal of Finance, 61*, 2551–2595.

Lee, C. F., Lee, A., & Liu, N. (2010). Alternative model to evaluate selectivity and timing performance of mutual fund managers: Theory and evidence. In J. B. Guerard Jr. (Ed.), *The*

handbook of portfolio construction: Contemporary applications of Markowitz techniques. New York: Springer.
Lee, C. F., & Rahman, S. (1990). Market timing, selectivity, and mutual fund performance: An empirical investigation. *Journal of Business, 63*, 261–278.
Lee, J. H., & Stefek, D. (2008). Do risk factors eat alphas? *Journal of Portfolio Management, 34*(4), 12–25.
Lintner, J. (1965a). The valuation of risk assets on the selection of risky investments in stock portfolios and capital investments. *The Review of Economics and Statistics,* 13–37.
Lintner, J. (1965b). Security prices, risk, and the maximum gain from diversification. *Journal of Finance,* 30.
Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.
Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly, 3*, 111–133.
Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investment*. New York: Wiley. Cowles Foundation Monograph No 16.
Markowitz, H. M. (1987). *Mean-variance analysis in portfolio choice and capital markets*. Oxford: Basil Blackwell.
Markowitz, H. M. (2000). *Mean-variance analysis in portfolio choice and capital markets*. New Hope, PA: Frank J. Fabozzi Associates.
Menchero, J., Morozov, A., & Shepard, P. (2010). Global equity risk modeling. In J. B. Guerard Jr. (Ed.), *The handbook of portfolio construction: Contemporary applications of Markowitz techniques*. New York: Springer.
Menchero, J., & Nagy, Z. (2015). Performance of earnings yield and momentum factors in US and international equity markets. In *Portfolio construction, measurement and efficiency*. New York: Springer.
Modigliani, F., & Miller, M. (1958). The cost of capital, corporate finance, and the theory of investment. *The American Economic Review, 48*, 261–297.
Modigliani, F., & Miller, M. (1963). Corporate income taxes and the cost of capital: A correction. *American Economic Review, 53*, 433–443.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*, 768–783.
Rosenberg, B. (1974). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis, 9*, 263–274.
Rosenberg, B., & Marathe, V. (1979). Tests of capital asset pricing hypotheses. In: Levy, H (Ed) *Research in finance* vol. 1.
Rosenberg, B., & McKibben, W. (1973). The prediction of systematic and unsystematic risk in common stocks. *Journal of Financial and Quantitative Analysis, 8*, 317–333.
Rosenberg, B., & Rudd, A. (1977). *The yield/beta/residual risk tradeoff*. Working paper no. 66, Research Program in Finance, Institute of Business and Economic Research, University of California, Berkeley.
Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341–360.
Ross, S. A., & Roll, R. (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance, 35*, 1071–1103.
Rudd, A., & Clasing, H. K., Jr. (1982). *Modern portfolio theory: The principles of investment management*. Homewood, IL: Dow Jones-Irwin.
Saxena, A., & Stubbs, R. A. (2012). An empirical case study of factor alignment using the user model. *Journal of Investing, 21*, 25–44.
Saxena, A., & Stubbs, R. A. (2015). Augmented risk models to mitigate factor alignment problems. *The Journal of Investment Management, 13*, 57–79.
Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science, 9*, 277–293.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 39*, 119–138.
Sharpe, W. F. (1970). *Portfolio theory and capital markets*. New York: McGraw-Hill.

Sivramakrishnan, K., & Stubbs, R. A. (2013). Improving the investment process with a custom risk model: A case study with the GLER model. *Journal of Investing, 22*, 129–144.

Stone, B. K. (1970). *Risk, return, and equilibrium: A general single-period theory of asset selection and capital market equilibrium*. Cambridge, MA: MIT Press.

Stone, B. K. (1974). Systematic interest-rate risk in a two-index model of returns. *Journal of Financial and Quantitative Analysis, 9*, 709–721.

Tobin, J. (1958). Liquidity preference as behavior to risk. *Review of Economic Studies, 25*, 65–86.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.

Treynor, J. L. (1999). Toward a theory of market value for risky assets. In R.A. Korajczyk (Ed.), *Asset pricing and portfolio performance*. London: Risk Books.

Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market. *Harvard Business Review, 44*, 131–136.

Treynor, J. (2008). *Treynor on institutional investing*. New York: Wiley.

Wei, K. D., Wermers, R. R., & Yao, T. (2016). Uncommon value: The investment performance of contrarian funds. In *Portfolio construction, measurement and efficiency*. New York: Springer.

Chapter 2
Portfolio Theory: Origins, Markowitz and CAPM Based Selection

Phoebus J. Dhrymes

The valuation of risky assets was initially based on bond valuation theory. Although the valuation of a bond may fluctuate due to variation in market interest rates, the coupon was fixed and subject mainly to the risk of default, which was episodic rather than continuous; prominent in the nature of the instrument were certain legal safeguards. When applied to stocks (risky assets) frequently the role of the coupon rate was played by the dividend, which though not fixed was deemed to be steady and subject only to infrequent changes. This framework, however, is evidently inappropriate in the case of stocks where the rate of return (principally earnings) is inherently variable and is not subject to legally binding specification.

The origin of modern finance in this context (portfolio selection) must be traced to the work of Markowitz (1952, 1956, 1959). Its basic framework is based on the work of von Neumann and Morgenstern (1944) (VNM) who pioneered the view that choice under uncertainty may be based on expected utility. The concept of utility is at least as old as the nineteenth century and the view that consumer choice (of the basket of goods and services consumed) was a compromise between the consumer's desires and the resources available to him (income). Thus, preceding expected utility constructs, the view prevailed that consumers obtained the most preferred bundle of goods and services they could attain with their incomes. But how could we import these concepts into the valuation of risky assets and their subsequent inclusion in a basket we call portfolio; after all consumers choose various goods because they satisfy some desire or group of desires. But a consumer (investor) need not have a preference or desire to own a given security per se. The importance of

I wish to thank John B. Guerard Jr. for helpful comments and stimulation.

P.J. Dhrymes (✉)
Columbia University, New York, NY, USA
e-mail: pdhrymes@gmail.com

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_2

Markowitz' contribution is that he isolated two aspects of relevance, return and risk, established a method of ranking them (a utility function), thus recognizing the inherent riskiness (randomness) of returns, and invoked VNM in the process. Having done so, it becomes clear that in this formulation the problem is conceptually broadly similar to the problem of consumer choice, although by no means identical. He correctly saw that it is not possible simultaneously to increase returns and at the same time minimize the risk entailed, because of arbitrage. Indeed, many of the later developments of the subject follow from these insights although not explicitly detailed in Markowitz (1959).

2.1 Constrained Optimization

Ignoring the utility or expected utility aspects, the (portfolio) selection problem was defined as: maximize expected returns subject to a variance and scale constraint.[1] Setting up the Lagrangian

$$\Lambda = \gamma' Er + \alpha r_0 + \lambda_1 (k - \gamma' \Sigma \gamma) + \lambda_2 (1 - e'\gamma - \alpha), \tag{2.1}$$

where E is the expectation operator, r is an n-element column vector containing the rates of return on the risky assets, r_0 is the risk free rate, $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)'$ is the portfolio composition, the individual elements γ_i denoting the proportion of the portfolio invested in the ith risky asset and α is the portion invested in the risk free asset; evidently, $\gamma' \Sigma \gamma$ is the variance of the portfolio, or its risk; it is assumed that at least for the duration of the choice period,

$$Er = \mu \qquad \mathrm{Cov}(r) = \Sigma > 0, \quad Er_0 = r_0 \quad \mathrm{var}(r_0) = 0. \tag{2.2}$$

If we solve for the first order conditions we find[2]

$$\gamma = \frac{1}{2\lambda_1} \Sigma^{-1} (\mu - er_0), \quad \alpha = 1 - \frac{1}{2\lambda_1} e' \Sigma^{-1} (\mu - er_0), \tag{2.3}$$

$$\lambda_1 = \frac{\gamma' (\mu - er_0)}{2\gamma' \Sigma \gamma}, \quad \lambda_2 = r_0, \quad e = (1, 1, 1, \dots 1)'. \tag{2.4}$$

[1]How does one explain that only the mean and variance of returns and not other moments play a role? One can justify this by an implicit assumption that the probability distribution of returns belongs to a family of distributions described by **only two** parameters, or that the expected utility function is of such a form that it depends only on the mean and variance of the relevant distribution.

[2]It should be noted that Markowitz did not actually solve for γ; rather his version focused only on risky assets and imposed **non-negativity constraints on the elements of** γ. Thus what he derived from the first order conditions were rules for inclusion in and/or exclusion from (of securities) in an optimal portfolio.

Although the solution was easy to obtain the interpretation of the Lagrange multiplier, λ_1 is clouded by the fact that **it is not invariant to scale**; thus if we were to double α and the elements of γ, the expression for the Lagrange multiplier would be halved without any change in other aspects of the procedure; thus any interpretation given to it in comparisons would be ambiguous and questionable. To that end we alter the statement of the constraint, thus redefining risk, to[3]

$$k = (\gamma' \Sigma \gamma)^{1/2} = \sigma_p.$$

without changing its substance. In turn this will yield the solution

$$\gamma = \frac{k}{\lambda_1} \Sigma^{-1}(\mu - er_0), \quad \alpha = 1 - \frac{k}{\lambda_1} e' \Sigma^{-1}(\mu - er_0), \tag{2.5}$$

$$\lambda_1 = \frac{\gamma'(\mu - er_0)}{\sigma_p}, \quad \lambda_2 = r_0, \quad e = (1, 1, 1, \ldots 1)'. \tag{2.6}$$

Examining the numerator of λ_1, i.e. **the Lagrange multiplier in the alternative formulation of the risk constraint** we find

$$\gamma'(\mu - er_0) = (\gamma' \mu + \alpha r_0) - r_0, \tag{2.7}$$

i.e. **it is the excess expected return of the portfolio** while the denominator is σ_p, i.e. the portfolio's risk! Thus the Lagrange multiplier attached to the risk constraint, in the Markowitz formulation, gives us the 'terms of trade' between reward and risk at the optimum. Noting further that

$$\frac{\partial \Lambda}{\partial k} = \lambda_1,$$

[3]From the point of view of computation, entering the constraint as $k^2 = \gamma' \Sigma \gamma$ simplifies operations, but makes the Lagrange multiplier harder to interpret in terms of common usage in finance; if, however, we enter the constraint as $k = (\gamma' \Sigma \gamma)^{1/2}$, we complicate the computations somewhat, we do not change the nature of the solution, but we can interpret the Lagrange multiplier in terms of common usage comfortably. We should also bear in mind that if risk is defined in terms of the standard deviation rather than the variance, a certain intuitive appeal is lost. For example, it is often said that security returns are subject to two risks, market risk and idiosyncratic risk. If we also say, as we typically do, that market risk is independent of idiosyncratic risk, then we have the following situation: denote the market risk by the variance of a certain random variable, say σ^2_{mar} and the idiosyncratic risk by the variance σ^2_{idio} then the risk of the security return is the **sum** $\sigma^2_{\text{mar}} + \sigma^2_{\text{idio}}$. On the other hand, if we **define risk in terms of the standard deviation**, then the two risks are not additive, i.e. the risk of the security is **not** $\sigma_{\text{mar}} + \sigma_{\text{idio}}$ but $\sqrt{\sigma^2_{\text{mar}} + \sigma^2_{\text{idi0}}}$, which is smaller, when we use as usual the positive square root. This problem occurs whenever there is aggregation of independent risks.

we may interpret λ_1 as the optimal marginal reward for risk or more correctly the marginal reward for risk at the optimum. All this is, of course, ex ante and assumes that the investor or the portfolio manager knows with certainty the mean and variance of the stochastic processes that determine ex post the realized returns.

2.2 Portfolio Selection and CAPM

Another aspect that needs to be considered is whether the index based on the interpretation of the Lagrange multiplier discussed in connection with the solution given to the portfolio selection model inv Markowitz (1959) is relevant in the CAPM context and whether these optimality procedures shed any light on the issue of composition rules.

For the latter issue, a more recent development along these lines is given in Elton et al. (2007), where the objective is stated as the maximization of the Sharpe ratio, which is the ratio of (expected) excess returns to (expected) standard deviation of a portfolio, using CAPM as the source of the covariance structure of the securities involved. It does that by means of nonlinear programming; from the first order conditions it derives rules of inclusion in (and exclusion from) an optimal portfolio. While similar in objective, this **is not** equivalent to the Markowitz approach. Moreover, it is questionable that maximizing the Sharpe ratio is an appropriate way for constructing portfolios. In particular, a portfolio consisting of a single near risk free asset with near zero (but positive) risk and a very small return might well dominate, in terms of the Sharpe ratio, any portfolio consisting of risky assets in the traditional sense. A ratio can be large if the numerator is large relative to the denominator, **or if the denominator is exceedingly small relative to a small positive numerator**. Consider (10/2) and (.5/0.1) or (.1/0.01). The point is that **given the level of risk** it is generally agreed that the higher the Sharpe ratio the better, however, to put it mildly, it is not generally accepted that the higher the Sharpe ratio the better, **irrespective of risk**. Evidently this would depend on the investor's or portfolio manager's tradeoff between risk and reward.

In Markowitz the rates of return are stochastic processes with fixed means and covariance matrix; thus what is being solved is an essentially static problem. It could be made somewhat dynamic by allowing these parameters to change over time, perhaps discontinuously.[4] This, however, imposes a considerable computational burden, viz. the re-computation of n means and $n(n+1)/2$ variances and covariances. On the other hand, if we adopt the framework of CAPM suggested, by Sharpe (1964), Lintner (1965b), Mossin (1966), Treynor (1962)[5] and others, as originally

[4]I say 'somewhat dynamic' because we still operate within what used to be called a 'certainty equivalent' environment, in that the underlying randomness is not fully embraced as in option price theory.

[5]The intellectual history of the evolution of CAPM is detailed in the excellent and comprehensive paper by French (2003), which details *inter alia* the important but largely unacknowledged role payed by the unpublished paper Treynor (1962). We cite Lintner (1965a) in the cite both Lintner paper of 1965 in his capital market development.

formulated, rates of returns are assumed to behave as

$$r_{ti} - r_{t0} = \beta_i(r_{mt} - r_{t0}) + u_{ti}, \quad i = 1, 2, \ldots, n. \quad t = 1, 2, \ldots T, \tag{2.8}$$

where r_{ti}, r_{t0}, r_{mt} are, respectively, the rates of return on the ith risky asset, the riskless asset and the market rate of return, β_i is a fixed parameter, at least in the context of the planning period; u_{ti} is, for each i, a sequence of independent identically distributed random variables with mean zero and variance ω_{ii}; moreover u_{ti} and $u_{t'j}$ are mutually independent for every pair (t, t') and (i, j). Notice that if we rewrite the CAPM equation as

$$r_{ti} = (1 - \beta_i)r_{t0} + \beta_i r_{mt} + u_{ti}, \tag{2.9}$$

this version of CAPM seems to assert that individual returns are, on the average, linear combinations (more accurately weighted averages for positive betas) of the risk free and market rates **with fixed weights**. A more popular recent version is

$$r_{ti} = c_i + \beta_i r_{mt} + u_{ti}, \tag{2.10}$$

where now c_i is an **unconstrained parameter**. If we bear in mind that the risk free rate is relatively constant it might appear that the two versions are equivalent. However, when considering applications this is decidedly not so. Some of the differences are

1. If we attempt to apply a (Markowitz) optimization procedure using the first version, the component α of the portfolio devoted to risk free assets cannot be determined and has to be provided *a priori*. This is due to the fact that in this version

$$Er_p = r_{t0} + \gamma' \beta(\mu_{mt} - r_{t0}),$$

which is the expected value of the returns on any portfolio $(\gamma,\ \alpha)$, **does not contain $\boldsymbol{\alpha}$** ; since the risk free rate has zero variance and zero covariances with the risky assets, α is not contained in the variance (variability) of the portfolio either. Thus, it cannot possibly be determined by the optimization procedure. With the alternative version, however, we can.
2. Bearing in mind that expected returns and risks are **not known and must be estimated prior to portfolio selection**, if we use the first version to determine an asset's beta we obtain

$$\hat{\beta}_i = \frac{\sum_{t=1}^{T}(r_{ti} - r_{t0})(r_{mt} - r_{t0})}{\sum_{t=1}^{T}(r_{mt} - r_{t0})^2}, \quad \hat{u}_{ti} = r_{ti} - \hat{\beta}_i(r_{mt} - r_{t0}), \qquad \hat{\omega}_{ii} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{ti}^2;$$

if we use the alternative (second) formulation of CAPM with an unrestricted constant term we would obtain

$$\tilde{\beta}_i = \frac{\sum_{t=1}^{T}(r_{ti} - \bar{r}_i)(r_{mt} - \bar{r}_m)}{\sum_{t=1}^{T}(r_{mt} - \bar{r}_m)^2}, \quad \tilde{u}_{ti} = r_{ti} - \hat{\beta}_i(r_{mt} - r_{t0}), \quad \tilde{\omega}_{ii} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{ti}^2;$$

If the risk free rate is appreciably smaller than the sample means of individual asset and market rates of return, the estimates of β_i could deviate appreciably from those obtained using the first version.

3. Equation (2.11) implies that every individual asset's rate of return is a linear combination of the risk free and market rates **but the coefficients of the linear combination need not be positive**. In particular it implies that an asset with negative beta does not respond to market rates as its beta might indicate, but the response is modulated by the term $(1 - \beta)r_{t0}$, which in this case is positive. In addition, it may have implications for well-diversified portfolios that have not yet been explored.

Thus, we shall conduct our analysis on the basis of the alternative (second) version of CAPM given in Eq. (2.12).

The main difference between our formulation and that in Markowitz is that here r_{mt} is a **random** variable with mean μ_{tm} and variance σ_{tm}^2 whose parameters **may vary with t, perhaps discontinuously**; it is, however, independent of $u_{t'i}$, for every pair (t, t') and i; moreover, if we use it as the basis for a Markowitz type procedure the resulting portfolios would depend on these parameters. Thus they could form the basis for explicit dynamic adjustment as their parameters vary in response to different phases in economic activity.

Within each t, the analysis is **conditional** on r_{mt}. The relation may be written, for a planning horizon T,

$$r_{ti} = c_i + \beta_i r_{mt} + u_{ti}, \quad i = 1, 2, \cdots, n \quad t = 1, 2, \cdots, T \tag{2.11}$$

where r_{ti}, r_{mt} are, respectively, the observations on the risk free and market rates at time t, c_i, β_i are parameters to be estimated and u_{ti} the random variables (error terms), often referred to as idiosyncratic risk, with mean zero and variance ω_{ii}. Because the analysis is done **conditionally** on r_{mt} and because **by assumption** the u_{ti} are independently distributed, and all equations contain the same (right hand, explanatory) variables, we can estimate the unknown parameters **one equation at a time without loss of efficiency, by means of least squares**. Now, can we formulate a Markowitz like approach in choosing portfolios on the basis of CAPM? Before we do so it is necessary to address an issue frequently mentioned in the literature, viz. that by diversification we may eliminate 'idiosyncratic risk'. What does that mean? It could simply mean that in a diversified portfolio idiosyncratic risk emanating from any one risky asset or a small class thereof is negligible relative to market risk, although it need not be zero. On the other hand, taken literally it means that

$$\lim_{n\to\infty}\sum_{i=1}^{n}\gamma_i\epsilon_{ti} \stackrel{\text{a.c.}}{\to} 0, \tag{2.12}$$

i.e. this entity converges to zero with probability 1, and thus idiosyncratic risk need not be taken into account, meaning that for the purpose of portfolio selection we can use a version of CAPM which **does not contain an idiosyncratic risk component**. Formally, what is required of such entity in order to converge (to its zero mean) with probability one? For example, in the special case where $\gamma_i \approx 1/n$, a sufficient condition for Eq. (2.14) to hold is given by Kolmogorov as[6]

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{\omega_{ii}}{i^2}\right) < \infty,$$

which would be satisfied if the ω_{ii} are bounded. For another selection of the components of γ it may not be; for example, if $\gamma_i \approx n^{\eta}/n, \eta > 0$ it will not be satisfied even if the variances are bounded. Since this assertion imposes a restriction on the vector, γ, of an undetermined nature, we prefer to explicitly take into account idiosyncratic risk in formulating the problem of optimal portfolio selection.

Another aspect that needs to be considered is whether the index based on the interpretation of the Lagrange multiplier discussed in connection with the solution given to the portfolio selection model in Markowitz (1959) is relevant in the CAPM context and whether these optimality procedures shed any light on the issue of composition rules.

We proceed basically as before except now the variability constraint utilizes the standard deviation. For clarity, we redefine portfolio returns and the covariance matrix of the securities involved given the CAPM specification; thus

$$r_p = \gamma' c + \gamma' \beta r_{mt} + \alpha r_{t0} + \gamma' u'_{t\cdot}, \quad \Sigma = \Omega + \sigma^2_{mt}\beta\beta', \tag{2.13}$$

and the solution is obtained by optimizing the Lagrangian

$$\Lambda = \gamma' c + \gamma' \beta r_{mt} + \alpha r_{t0} + \lambda_1[k - (\gamma'\Sigma\gamma)^{1/2}] + \lambda_2(1 - \gamma' e - \alpha), \tag{2.14}$$

From the first order conditions we easily obtain

$$\left(\Omega + \sigma^2_{mt}\beta\beta'\right)\gamma = \frac{k}{\lambda_1}\left(c + \beta\mu_{tm} - e r_{t0}\right) \tag{2.15}$$

$$\alpha = 1 - \gamma' e, \quad \lambda_2 = r_{t0}, \quad e = (1, 1, \ldots 1)' \tag{2.16}$$

$$\lambda_1 = \frac{E r_p - r_{t0}}{[\gamma'(\Omega + \sigma_t^2 \beta\beta')^{\gamma}]^{1/2}}. \tag{2.17}$$

[6]See Dhrymes (2013, pp. 202–203).

The last equation is easily obtained by premultiplying the first equation above by γ' and using the definition of Er_p implied by Eq. (2.15) above. If we now substitute for λ_1 we obtain an equation that involves only α and γ, i.e.

$$(\gamma' c + \gamma' \beta \mu_{mt} + \alpha r_{t0})\gamma = k^2 \left(\Omega + \sigma_{mt}^2 \beta\beta'\right)^{-1} (c + \beta\mu_{mt} - er_{t0}). \qquad (2.18)$$

But, if we use Eq. (2.16) we can eliminate α so that Eq. (2.18) may be rewritten as

$$\gamma\gamma'(c + \beta\mu_{mt} - er_{t0}) - \gamma r_{t0} = k^2 \left(\Omega + \sigma_{mt}^2 \beta\beta'\right)^{-1} (c + \beta\mu_{mt} - er_{t0}), \qquad (2.19)$$

which can now be solved for γ.

A number of features of this procedure need to be pointed out:

1. No high dimensional matrix needs to be inverted, due to a result (Corollary 2.5),[7] which enables us to write

$$\left(\Omega + \sigma_{mt}^2 \beta\beta'\right)^{-1} = \Omega^{-1} - \zeta\Omega^{-1}\beta\beta'\Omega^{-1}, \quad \zeta = \frac{\sigma_{mt}^2}{1 + \sigma_{mt}^2 \beta'\Omega^{-1}\beta};$$

   since Ω **is diagonal** we easily compute

$$\beta'\Omega^{-1}\beta = \sum_{i=1}^{n} \left(\frac{\beta_i^2}{\omega_{ii}}\right), \quad \Omega^{-1}\beta\beta'\Omega^{-1} = \left[\frac{\beta_i\beta_j}{\omega_{ii}^2}\right],$$

   i.e., it is a matrix whose typical element is $\beta_i\beta_j/\omega_{ii}^2$.
2. The number of parameters that we need to estimate prior to optimization is $3n+2$, viz. the elements of the vectors c, β and the variances ω_{ii}; all of these can be obtained from the output of n simple regressions. The other two parameters are simply the mean and variance of the market rate.
3. The procedure yields a set of equations which are quadratic in γ; the solution is a function of k^2, μ_{mt}, σ_{mt}^2 and can be adjusted relatively easily when updating of the estimates of μ_{mt}, σ_{mt}^2 is deemed appropriate.
4. It is interesting that the optimal (solution vector) composition vector, γ, is a **function of (depends on) the risk parameter** k^2, **not** k, i.e. **risk is represented by the variance, not the standard deviation**.

We thus see that in the context of CAPM the implementation of optimal portfolio selection becomes much simpler and computationally more manageable and, consequently, so is the task of evaluation *ex post*.

[7]See Dhrymes (2013, pp. 46–47).

2.3 Conclusion

In this paper we reconsidered the problem of portfolio selection as formulated by Markowitz (1959) and proposed an extension based on CAPM. This extension highlights certain aspects that represent a considerable simplification; it illuminates issues regarding the estimation of securities betas, the role played by idiosyncratic risk and leads to the formulation of a set of quadratic equations that define the optimal composition of efficient portfolios (the elements of the vector γ), as a function of the selected level of risk and estimates of (expected) market rate and its risk (variance). The only remaining problem is to find an algorithm that solves sets of quadratic equations. This should not be very difficult. Given that, it offers a systematic way in which portfolio managers might insert into the process their evolving views of market rates and their associated risk, when updating is deemed necessary.

An interesting by-product is the potential provided by this framework in evaluating (managed) portfolio performance. In a now classic paper Sharpe (1966) evaluates mutual fund performance by considering realized rates of return for a number of mutual funds over a number of years and computes the standard deviation of such returns. The evaluation relies on the ratio of average returns to their standard deviation. Strictly speaking, these two measures do not estimate 'constant parameters' since the composition of the fund is likely to have changed appreciably over the period; thus their ratio is not a ranking of the fund itself. It is, however, **a ranking of the fund cum manager**.

If we use the framework presented in the paper which is based on CAPM we could, in principle, during each period compute from published data the portfolio or fund risk as $\gamma'(\Omega + \sigma_m^2 \beta\beta')\gamma$. Thus, the evaluator will have for each period, **both realized returns and risk**. This would make a more satisfactory basis for evaluation.

References

Dhrymes, P. J. (2013). *Mathematics for econometrics* (4th ed). Berlin: Springer.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzman, W. N. (2007). *Modern portfolio theory and investment analysis* (7th ed.). New York: Wiley.

French, C. (2003). The Treynor capital asset pricing model. *Journal of Investment Management, 1*, 60–72.

Lintner, J. (1965a). The valuation of risk assets on the selection of risky investments in stock portfolios and capital investments. *The Review of Economics and Statistics, 47*, 13–37.

Lintner, J. (1965b). Security prices, risk, and the maximum gain from diversification. *Journal of Finance, 30*, 587–615.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly, 3*, 111–133.

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investment*. Cowles Foundation Monograph (Vol. 16). New York: Wiley.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*, 768–783.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business: A Supplement, 1*(2), 119–138.
Treynor, J. L. (1962). Toward a theory of market value of risky assets. Later published in R. A. Korajczyk (Ed.), *Asset Pricing and Portfolio Performance* (pp. 15–22). London: Risk Books; Unpublished manuscript, rough draft dated by J.L. Treynor Fall 1962.
von Neumann, J., & Oskar M. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Chapter 3
Market Timing

Ravi Jagannathan and Robert A. Korajczyk

A large fraction of retirement savings, valued at $23 trillion at the end of March 2014 in the USA, is managed by professional money managers (Investment Company Institute 2014). For example, mutual funds accounted for 60 % of households' assets in defined contribution retirement plans and 45 % of the assets in individual retirement account plans. Investment companies and pension funds held almost 40 % of total corporate equities in the first quarter of 2014 (Board of Governors of the Federal Reserve System 2014). While investors benefit by delegating the money management function to professional managers with specialized talents and skills, relying on others creates invisible indirect costs in addition to observable direct fees that those managers charge. These invisible costs arise from the need to monitor and evaluate the actions of the managers in order to ensure that those actions are consistent with investors' objectives and the explicit and implicit contractual terms. An investor allocating savings across several managers has to ensure that the investment objectives are satisfied, the bets that the different managers take do not cancel out, and the funds allocated to a particular manager are consistent with the manager's abilities and investment capacity. Hence, there is a need for a conceptual framework for evaluating and monitoring delegated fund managers.

We thank Wayne Ferson, John Guerard, Winfried Hallerbach, Maja Kos, and Selwyn Yuen for helpful comments.

R. Jagannathan • R.A. Korajczyk (✉)
Kellogg School of Management, Northwestern University, 2001 Sheridan Road,
Evanston, IL 60208-2001, USA
e-mail: rjaganna@kellogg.northwestern.edu; r-korajczyk@kellogg.northwestern.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_3

An investor always has the safe-harbor default option of allocating the savings across judiciously chosen passively managed index funds, each of which represents an asset class.[1] Therefore, it is natural to start with an appropriately chosen portfolio of such index funds as the benchmark for evaluating the performance of an active portfolio manager. In order to minimize the hidden and unknown risks that arise from delegating the portfolio management function, a fund manager's mandate will often limit the investment universe to securities that are in the benchmark and other securities that have certain characteristics. The weights assigned to the various securities in the manager's active portfolio will necessarily have to deviate from their weights in the benchmark if superior performance is to be delivered. The magnitude of such deviations will be limited by mandate restrictions.

The common practice is to decompose a fund manager's performance relative to the benchmark into two components: (a) the ability to identify which asset classes will perform better (and overweight those asset classes relative to the benchmark) and (b) the ability to identify securities within each asset class that will do better than others in the same asset class. The former skill is denoted as "allocation" or "timing" and the latter as "selection." Such a decomposition helps in at least three ways. First, it facilitates assessing the extent to which past performance may carry over into the future by providing a better understanding of the nature of a manager's skill set and, therefore, the predictability of performance. Second, as we will see later, a manager with timing skill provides the investor with valuable portfolio insurance. Valuing such insurance features involves the use of contingent claims valuation methods, and the standard CAPM or linear beta pricing models will typically understate the value of such features. Third, attempts to add value through security selection and timing expose investors to different sources of risk. An appropriate decomposition assists the investor in understanding the exposure attained through a managed portfolio.

When portfolio holdings are observed, attributing the ex post performance to these two components is straightforward. However, for many investors, managed portfolio holdings are reported infrequently, typically at quarterly intervals, yet they are subject to continuous revisions. Further, assessing whether the manager's skills are significant enough to consistently outperform the benchmark in the future requires a theoretical framework. For example, a manager might attempt to exploit the low-frequency nature of return or holdings observation. The theoretical underpinnings of such a strategy allow one to design tests for its existence. Our focus here is on assessing the performance of a fund manager based on the portfolio's historical returns and inferring the sources of superior performance, if any.

[1] The question of how to choose from the vast collection of index funds to construct a portfolio to meet one's investment objectives has not received much attention in the academic literature and has been left to journals catering to the needs of investment advisers and individual and institutional investors. The exceptions include Elton et al. (2004) and Viceira (2009).

When the benchmark has only two asset classes, the aggregate market portfolio and the risk-free asset, allocation is referred to as "market timing." Treynor and Mazuy (1966) developed the basic theoretical framework for inferring the market-timing skills of a portfolio manager based on observations of the returns the manager generates under the assumption that the return on the market portfolio is unpredictable. Pfleiderer and Bhattacharya (1983) and Admati et al. (1986) examine the information structure and objectives of fund managers when the Treynor and Mazuy (1966) quadratic-regression approach can be used to identify the factor-timing skills of a portfolio manager. Merton (1981) developed the framework for assessing the value added by a portfolio manager who can successfully time the market. Henriksson and Merton (1981) empirically evaluate the performance of mutual fund managers using the Merton (1981) framework to separate their selection and timing skills. Subsequent researchers have relaxed the assumptions along several dimensions: Henriksson and Merton (1981) allow for timing across several asset classes; Jagannathan and Korajczyk (1986) expand the universe of securities to include assets that have embedded passive market timing-like features; Glosten and Jagannathan (1994) allow more-general trading and portfolio-rebalancing behavior on the part of the manager leading to complex option-like features in managed portfolio returns and show how to assess the value added by such a manager; and Ferson and Schadt (1996) allow predictability of the market portfolio return, and funds' market exposures, based on publicly available information. Ferson and Mo (2016) show how market return and volatility timing, discussed in Busse (1999), can be separated when portfolio weights are observed.

The literature that has grown out of the work of Treynor and Mazuy (1966) allows us to address several critical questions. Is there be a meaningful distinction between forecasting security-specific returns and forecasting systematic factor returns, particularly in a world with dynamic trading strategies and portfolio containing derivative securities? Do standard performance measures give accurate indications of the sum of selection and timing performance? If the market risk premium varies through time in predictable ways, how do we distinguish between timing based on public information versus timing based on true skill?

In Sect. 3.1 we discuss the main focus of this chapter, return-based performance measurement. We begin with measures that do not explicitly incorporate timing, the Treynor (1965) and Jensen (1968) measures. These are related to measures that test for timing ability, particularly the quadratic Treynor and Mazuy (1966) measure and piecewise-linear Henriksson and Merton (1981) measure. We then discuss the effects of derivative strategies, dynamic trading strategies at higher frequencies than return observations, and pseudo timing, on portfolio performance evaluation. The section finishes with a discussion of performance evaluation when risk premia have predictable components. Section 3.2 contains a brief discussion of holdings-based performance measures and Sect. 3.3 concludes.

3.1 Return-Based Performance Measurement

Modern asset pricing models make the important distinction between the systematic and idiosyncratic components of asset returns. The former is correlated with investors' marginal utility and, therefore, commands a risk premium, while the latter has no risk premium. This naturally leads to portfolio performance metrics that decompose portfolio returns into a component due to exposure to systematic risk and a component due to the ability to forecast the idiosyncratic returns of assets. The component due to exposure to systematic risk can be obtained through passive portfolios while the component due to the ability to forecast the idiosyncratic returns of assets (which are unconditionally unforecastable) represents skill. The skill component is often referred to as the risk-adjusted return, the abnormal return, Jensen measure, or alpha of the portfolio. This decomposition—or a scaling of it—is originally proposed in Treynor (1965) and Jensen (1968).

In the context of the single-index capital asset pricing model (CAPM) of Treynor (1962, 1999), Sharpe (1964), Lintner (1965), and Mossin (1966), systematic risk is measured by the sensitivity of asset returns to unexpected returns on the aggregate market portfolio. Let $r_{i,t}$ denote the return on asset i in period t; $r_{f,t}$ denote the return on a riskless asset in period t; $R_{i,t}$ denote $r_{i,t} - r_{f,t}$; $R_{m,t}$ denote the excess return on the aggregate market portfolio; and $\delta_{m,t}$ denote $R_{m,t} - E[R_{m,t}]$. The return generating process for asset returns is assumed to be

$$R_{i,t} = E[R_{i,t}] + \beta_{i,m}\delta_{m,t} + u_{i,t}. \tag{3.1}$$

Unexpected asset returns are driven by shocks to the market portfolio, $\delta_{m,t}$, and shocks uncorrelated with the market, $\varepsilon_{i,t}$. Under the assumptions of the CAPM, investors marginal utility is perfectly correlated with returns on the market. Therefore, investors demand higher returns on assets with greater exposure to market risk, and expected excess returns on assets are determined by $\beta_{i,m}$, in equilibrium:

$$E[R_{i,t}] = \beta_{i,m}E[R_{m,t}]. \tag{3.2}$$

Combining the data generating process (3.1) and the equilibrium model (3.2), we obtain

$$R_{i,t} = \beta_{i,m}[E[R_{m,t}] + \delta_{m,t}] + u_{i,t}, \tag{3.3}$$

or, equivalently,

$$R_{i,t} = \beta_{i,m}R_{m,t} + u_{i,t}. \tag{3.4}$$

When a portfolio manager possesses superior skills in evaluating individual assets, that manager's expected value of the nonsystematic returns on assets will be nonzero, thus leading to a relation, conditional on the manager's private signal,

$$R_{i,t} = \alpha_{i,t} + \beta_{i,m}R_{m,t} + \varepsilon_{i,t} \tag{3.5}$$

where $\alpha_{i,t} = E[u_{i,t}|X_{t-1}]$ and X_{t-1} is the information set of the informed manager. For an investor without skill, the unconditional expectation is zero: $E[\alpha_{i,t}] = 0$.

A portfolio manager with such skills would take active positions in assets with nonzero values of $\alpha_{i,t}$. That is, the manager will overweight assets, relative to their market weights, with $\alpha_{i,t} > 0$ and underweight assets with $\alpha_{i,t} < 0$ as in Treynor and Black (1973). Therefore, the portfolio-return generating process ($R_{p,t}$) for a manager with skill would be

$$R_{p,t} = \alpha_p + \beta_{p,m} R_{m,t} + \varepsilon_{p,t}, \tag{3.6}$$

with $\alpha_p > 0$. This risk-adjusted performance metric is first suggested in Treynor (1965) and Jensen (1968) and is commonly called the Jensen measure of performance. Treynor (1965) notes that α_p is a function of the manager's skill in predicting $\varepsilon_{i,t}$ and the aggressiveness with which the manager uses that information. For example, two managers with the same risky asset portfolio, but different levels of leverage, will produce different Jensen measures. Consider a manager who constructs portfolio p, with performance α_p, and a manager who takes 50 % leverage (at the riskless rate) and invests the proceeds in an identical portfolio. Call this levered portfolio q. $R_{q,t} = 1.5 \times R_{p,t}$, which implies $\alpha_{q,t} = 1.5 \times \alpha_{p,t}$ and $\beta_{q,t} = 1.5 \times \beta_{p,t}$. In this case, the manager of portfolio q has no more skill than the manager of portfolio p. The higher value of alpha is merely a reflection of manager q's higher level of aggressiveness. Treynor (1965) suggests a variant of the Jensen measure that controls for the aggressiveness of the manager. Treynor's measure, which we denote T_p, scales Jensen's measure by the beta of the portfolio,

$$T_p = \frac{\alpha_p}{\beta_{p,m}}. \tag{3.7}$$

With this scaling $T_q = \frac{1.5 \times \alpha_p}{1.5 \times \beta_{p,m}} = \frac{\alpha_p}{\beta_{p,m}} = T_p$, giving a cleaner measure of skill adjusted for aggressiveness. The Jensen and Treynor measures extend naturally to other asset pricing models, such as the multifactor models of Merton (1973b) and Ross (1976), and the use of the single-index CAPM is merely meant to simplify the exposition.

Treynor and Black (1973) developed the above framework further, showing how the information in Eq. (3.7) can be used by an investor, who maximizes the expected return on her wealth portfolio subject to a target variance constraint, to allocate her wealth across actively managed funds and the market portfolio. Suppose the market portfolio has an expected return $E[R_m]$, variance σ_m^2, there is a single actively managed portfolio A with Jensen's alpha, α_A, and variance of the residual return in Eq. (3.6), $\sigma_{\varepsilon A}^2$. Then the investor who chooses a mean-variance efficient portfolio of the market portfolio and portfolio A will have a weight w_A in A and a weight $(1-w_A)$ in the market portfolio, where

$$w_A = \frac{\left[\alpha_A / \sigma_{\varepsilon A}^2\right]}{\left[E(R_m)/\sigma_m^2\right]}.$$

An investor will in general have access to several actively managed funds, $i, i = 1, 2 \ldots N$. Treynor and Black (1973) show that the weight w_i of the i'th actively managed fund in the portfolio of actively managed funds, A, will be given by

$$w_i = \sum_{k=1}^{N} \frac{\left[\alpha_i / \sigma_{\varepsilon i}^2\right]}{\left[\alpha_k / \sigma_{\varepsilon k}^2\right]}. \tag{3.8}$$

In deriving the above results, Treynor and Black (1973) assumed that $\mathrm{Cov}(\varepsilon_i, \varepsilon_k) = 0$. With this assumption, an investor can calculate how much to allocate across actively managed funds and the market portfolio based on knowledge of the parameters in Eq. (3.6) using a hand-held calculator. With the advent of personal computers, and spreadsheet software, the use of portfolio optimization has become prevalent among institutional investors and financial advisers. While the underlying framework is the same as in Treynor and Black (1973), it is no longer necessary to assume that $\mathrm{Cov}(\varepsilon_i, \varepsilon_k) = 0$, and multifactor extensions of Eq. (3.8) are widely used. However, the Treynor and Black (1973) model provides a huge amount of insight and intuition. It is still the textbook example for expositing the general principles underlying optimally combining actively managed funds with the market portfolio.[2]

In the formulation above, it is assumed that the manager's skill is in predicting the asset-specific, nonsystematic component of returns, $\varepsilon_{i,t}$. While there are many investors for whom this is an accurate description of their investment strategy, there are many other investors whose explicit strategy is to forecast market, or asset class, returns and adjust exposures to systematic risk (e.g., $\beta_{p,m}$ in the single-index, CAPM context) to take into account those forecasts. Such strategies have come under various labels, including market timing, tactical asset allocation, global macro investing, and others. In the analysis of Treynor and Black (1973), the optimal overall market exposure is determined by the market Sharpe Ratio expected by the investor, $\frac{E(R_{m,t})}{\sigma_m}$, scaled by $\frac{1}{\sigma_m}$. Therefore, the portfolio beta would vary with the expected market Sharpe Ratio, consistent with market timing.

As we mentioned earlier, managers who can successfully time the market provide portfolio insurance, and the standard mean-variance optimization framework is inadequate in evaluating such fund managers and deciding how much to allocate to them. Further, the potential existence of market-timing skills raises a number of important issues for performance evaluation, including:

1. Can there be a meaningful distinction between forecasting security-specific returns and forecasting systematic factor returns? For example, in the CAPM example used above, the market return is a market-capitalization weighted average of the individual asset returns,

[2] For example, see Bodie et al. (2011, Chap. 27.1).

$$R_{m,t} = \sum_{i=1}^{N} \omega_{j,t-1} R_{i,t}, \tag{3.9}$$

so having forecasting skills for $u_{i,t}$ must tell you something about $R_{m,t}$ since it contains a linear combination of the realizations of $u_{i,t}, i = 1, 2, \ldots n$. One approach is to think of $\alpha_{i,t}$ as the return implied by the manager's information, X_{t-1}, after accounting for any implications for $R_{m,t}$ [see, Admati et al. (1986, Sect. I)]. This implies that X_{t-1} cannot contain information only about $\alpha_{i,t}$ since the weighted average alpha most be zero. Another approach is to assume that asset returns are driven by an underlying factor model [see, Admati et al. (1986, Sect. I)].

2. Will the Jensen and Treynor measures give accurate indications of the sum of performance due to micro forecasting skills (often referred to as security selection) and macro forecasting skills (often referred to as market timing or asset allocation) of the portfolio manager?
3. Given the existence of securities and dynamic trading strategies that yield payoffs that are nonlinear in market (or factor) returns, can one separately measure the performance due to security selection and market timing, and does this dichotomy make sense?
4. Is it possible to create pseudo-timing performance? If so, how would that manifest itself in asset returns and performance measures?
5. If the market risk premium varies through time in predictable ways, how do we distinguish between timing based on public information versus timing based on true skill? For example, consider the following decomposition, where $E[R_m]$ is the unconditional market risk premium:

$$\Delta_{m,t} = R_{m,t} - E[R_m] = [R_{m,t} - E[R_{m,t}]] + [E[R_{m,t}] - E[R_m]] = \delta_{m,t} + \delta^*_{m,t}. \tag{3.10}$$

True market-timing ability is the ability to predict $R_{m,t}$ over and above the market's conditional expectation, $E[R_{m,t}]$, i.e., the ability to predict $\delta_{m,t}$. However, having ability to predict $\delta^*_{m,t}$ reflects one's ability to measure changes in the market's conditional risk premium.

3.1.1 Treynor and Mazuy (1966)

The pioneering paper in the measurement of market-timing ability is Treynor and Mazuy (1966). The essence of market timing or tactical asset allocation is to increase the portfolio's exposure to the market or a particular asset class when the manager expects high returns in that asset class and to decrease the portfolio's exposure when the manager expects low returns. When the manager has ability to forecast $\delta_{m,t}$ or uses public information to predict $\delta^*_{m,t}$, there will be a convex relation between the return on the portfolio, $R_{p,t}$, and the return on the market or asset class, $R_{m,t}$. Treynor and Mazuy propose that the convex relation be approximated by a quadratic relation,

$$R_{p,t} = a_p + b_{p,m}R_{m,t} + c_{p,m}R_{m,t}^2 + \varepsilon_{p,t}. \tag{3.11}$$

In this specification, $c_{p,m} > 0$ would be consistent with shifting into high-exposure assets when the manager's conditional expectation of $\delta_{m,t}$ is high or when $E[\delta_{m,t}^*]$ is high. Ability to forecast individual-asset nonsystematic returns, $u_{i,t}$ in (3.4), would, presumably, be captured by a_p (in the models of the return generating process and manager behavior in Jensen (1972) and Pfleiderer and Bhattacharya (1983) discussed below). That is, $a_p = \alpha_p^S$, where α_p^S is the alpha generated by *security selection* ability, while timing ability is captured by $c_{p,m}$.

Treynor and Mazuy (1966) apply their measures to 57 mutual funds over a 10-year period. The requirement of funds having a complete 10 years of data probably imparts an upward bias to any performance measures. Even so, only one of the 57 funds has a significantly positive value of $\hat{c}_{p,m}$ at the 5 % level of significance. Just by chance, one would expect that three funds would show significantly positive values of $\hat{c}_{p,m}$ since 5 % of 57 is 2.85 (assuming $\varepsilon_{p,t}$ is independent across funds). For this one significant fund, the positive value of $\hat{c}_{p,m}$ is accompanied by negative security selection ($\hat{\alpha}_p^S < 0$). We will address the negative cross-sectional relation between $\hat{c}_{p,m}$ and $\hat{\alpha}_p^S$ later.

Treynor and Mazuy conclude that there is little evidence to support the existence of timing ability in the sample of mutual funds they study.

3.1.2 The Relation Between $\beta_{p,m}$ and $R_{m,t}$

The quadratic functional form in (3.11) proposed by Treynor and Mazuy (1966) is meant to capture the notion that timing ability should exhibit a positive relation between market exposure, $\beta_{p,m}$, and $R_{m,t}$, which results in a convex relation between $R_{p,t}$ and $R_{m,t}$. A second-order polynomial is one way to approximate any general convex relation. The actual relation between $R_{p,t}$ and $R_{m,t}$ would be determined by the manner in which portfolio managers utilize any forecasting ability that they have.

3.1.2.1 Quadratic Characteristic Line

The quadratic relation is implied by reasonable models of manager behavior, studied in Jensen (1972), Pfleiderer and Bhattacharya (1983), and Admati et al. (1986). Assume that $\Delta_{m,t}$ is conditionally normally distributed and that the manager receives a signal, s_{t-1}, about $\Delta_{m,t}$ with $s_{t-1} = \Delta_{m,t} + \eta_t$, with η_t being a zero-mean normally distributed random variable independent of $\Delta_{m,t}$ (Jensen (1972) also allows for biased signals). $E[\Delta_{m,t}|s_{t-1}] = \lambda \times s_{t-1}$ with $\lambda = \frac{\sigma_\Delta^2}{\sigma_\Delta^2 + \sigma_\eta^2}$.

For quadratic utility (Jensen 1972) or constant absolute risk tolerance (Admati et al. 1986; Pfleiderer and Bhattacharya 1983), the market exposure of the optimal portfolio is linear in s_{t-1}:

$$\beta_{p,m,t} = \beta_{p,m} + \theta \times \lambda \times s_{t-1}, \tag{3.12}$$

where $\theta = \frac{1}{a\text{Var}(\Delta_{m,t}|s_{t-1})}$ and a is the risk aversion that the manager assumes for the fund investors. The unconditional average beta is $\beta_{p,m}$, and the period-by-period betas deviate from $\beta_{p,m}$ depending on the manager's signal, s_{t-1}. The aggressiveness with which the manager adjusts market exposure depends on risk aversion and the quality of the signal the manager receives about $\Delta_{m,t}$, through the influence of $\text{Var}(\Delta_{m,t}|s_{t-1})$ on θ and through λ. Additionally, the target level of beta (when $s_{t-1} = 0$) is given by $\theta E[R_m]$ (Pfleiderer and Bhattacharya 1983, p. 8). This leads to the following data generating process for portfolio returns:

$$R_{p,t} = \alpha_p^S + \theta E[R_m](1-\lambda)R_{m,t} + \theta\lambda R_{m,t}^2 + u_{p,t}, \tag{3.13}$$

where

$$u_{p,t} = \theta\lambda R_{m,t}\eta_t + \varepsilon_{p,t}.$$

The expected value of $u_{p,t}$ is zero since $R_{m,t}$ and η_t are independent. However, the residuals in (3.13) exhibit conditional heteroskedasticity. From (3.13) we see that the parameters in the regression of Treynor and Mazuy (1966) are given by

$$a_p = \alpha_p^S \tag{3.14}$$

$$b_{p,m} = (1-\lambda)\beta_{p,m} \tag{3.15}$$

$$c_{p,m} = \theta\lambda. \tag{3.16}$$

Thus, in this setting the Treynor and Mazuy (1966) intercept (a_p) and coefficient on the quadratic term ($c_{p,m}$) are consistent estimates of security-selection (α_p^S) and market-timing ($\theta\lambda$) skills. However, the coefficient on the linear term ($b_{p,m}$) is a downward-biased estimate of the target beta when the manager has ability to forecast the market return ($\lambda > 0$).

When the manager has both security-selection and market-timing skills and follows the Pfleiderer and Bhattacharya (1983) investment strategy, the fund data generating process follows (3.13). One could estimate the Jensen and Treynor measures specified in (3.6) and (3.7) for the fund. A reasonable assumption is that the estimated Jensen measure, $\hat{\alpha}_p$, would reflect both security selection, α_p^S, *and* the fact that the fund earns a higher return than one would expect given its average market beta, due to market-timing skill. In fact, this is not necessarily true since the unconditional fund beta in (3.6) yields a biased estimator of the average beta, $\theta E[R_m]$ (Dybvig and Ross 1985; Grant 1977; Jensen 1968, 1972). Using (3.13) and the definition of the unconditional beta, we can derive the bias under the assumed manager behavior:

$$E(\hat{\beta}_{p,m}) = \frac{\mathrm{Cov}(R_{p,t}, R_{m,t})}{\sigma_m^2} = \theta(1 + \lambda)E[R_m] + \theta\lambda\sigma_m Sk_m, \tag{3.17}$$

so that

$$E(\hat{\beta}_{p,m}) - E(\beta_{p,m,t}) = \theta(1+\lambda)E[R_m] + \theta\lambda\sigma_m Sk_m - \theta E[R_m] = \theta\lambda E[R_m] + \theta\lambda\sigma_m Sk_m,$$

where $Sk_m = \frac{E(R_{m,t} - E[R_m])^3}{\sigma_m^3}$ is the coefficient of skewness of the market return. Under the assumption of normality of market returns, $Sk_m = 0$ and the last term drops out of the expression. When the manager has no market-timing skills ($\lambda = 0$), $E(\hat{\beta}_{p,m}) = E(\beta_{p,m,t})$. From (3.17) we can determine the expected value of Jensen's performance measure:

$$E(\hat{\alpha}_p) = \alpha_p^S + \theta\lambda(\sigma_m^2 - E[R_m]^2) - \theta\lambda\sigma_m E[R_m]Sk_m. \tag{3.18}$$

$E(\hat{\alpha}_p)$ clearly reflects security-selection skill, α_p^S, but $E(\hat{\alpha}_p)$ could be either higher or lower than α_p^S even when the manager has timing skill. A manager with timing skill creates a portfolio with non-normal returns even in a world where primitive assets have normally distributed returns. This is due to the fact that portfolio returns include terms that are the product of the manager's normally distributed signal and normally distributed returns, leading to the quadratic term in (3.13). The linear specification in Jensen's alpha does not take into account the skewness induced by the manager's skill.

A simple numerical example may be useful here. Let us assume that $E[R_m] = 0.10, \sigma_m = 0.20$, and the market returns are normal. In this case,

$$E(\hat{\alpha}_p) = \alpha_p^S + \theta\lambda \times 0.03,$$

so that Jensen's measure reflects both security selection and timing ability, although the measured timing ability is likely to be biased. We will return to this issue later. However, other parameter specifications and market skewness can lead to a Jensen's measure that is either above, or below, α_p^S. When there is no timing ability ($\lambda = 0$), Jensen's measure provides an unbiased measure of α_p^S, and this is true regardless of whether the manager reacts optimally to the signal (Jensen 1972). In this case, Treynor's measure, T_p, is also consistent.

3.1.2.2 Piecewise-Linear Characteristic Line

The quadratic relation between portfolio returns and market returns is consistent with a world in which managers receive noisy, normally distributed signals about future market returns and behave as if they maximize a constant absolute risk tolerance utility function. An alternate assumption about manager behavior is that managers receive a signal about whether the market excess return will be positive. They then choose between two levels of exposure to systematic risk: high beta when they expect positive excess returns and low beta if they expect negative excess returns. Merton (1981) shows that timing ability in this setting is equivalent to the manager creating free call options on the market index. Through put–call parity, timing ability is also equivalent to creating a free protective put strategy. Therefore, the value created by the timing ability is given by the value of the number of free options created by timing skill (less the manager's fee).

Henriksson and Merton (1981) develop both nonparametric and parametric methods for evaluating timing and security-selection skills. Under the assumed manager behavior, the data generating process for portfolio returns is

$$R_{p,t} = a_p^S + \beta_{p,m}^U R_{m,t} + \beta_{p,m}^{U-D} \max[0, -R_{m,t}] + \varepsilon_{p,t}. \tag{3.19}$$

In (3.19) a_p^S measures security-selection skill (under the assumed return generating process and managerial behavior), $\beta_{p,m}^U$ measures the beta of the portfolio during "up" markets (markets where $R_{m,t} > 0$), and $\beta_{p,m}^{U-D}$ measures the difference between the portfolio's beta in "up" markets and its beta in "down" markets. Successful timing skill should result in a positive value of $\beta_{p,m}^{U-D}$. In the option-based framework, $\beta_{p,m}^{U-D}$ is the number of free call options on the market generated by the manager's skill at timing.

Henriksson (1984) estimates (3.19) for a sample of 116 open ended mutual funds. The average of the estimated values of $\hat{\beta}_{p,m}^{U-D}$ is negative, and 62 % of the funds studied have negative values of $\hat{\beta}_{p,m}^{U-D}$, consistent with the findings of Treynor and Mazuy (1966). This seemingly anomalous evidence of negative timing skill (present whether the quadratic or piecewise-linear specification is estimated) has proven to be remarkably robust and is observed for both mutual funds (Henriksson 1984; Jagannathan and Korajczyk 1986; Kon 1983), bond funds (Chen et al. 2010), and hedge funds (Asness et al. 2001; French and Ko 2006; Connor et al. 2010, Chap. 13), although not universally observed (Chen and Liang 2007). Hallerbach (2014) discusses the difference between the quadratic and piecewise-linear specifications for a portfolio's information ratio (defined by the conditional expected active return divided by active risk).

3.1.3 Derivative Strategies, Frequent Trading, Pseudo Timing, and Portfolio Performance

3.1.3.1 Derivative Strategies and Pseudo Timing

In evaluating the performance of a portfolio manager based only on observations of historical returns on a manager's portfolio, we rely on the assumption that the return on any primitive asset i is generated according to Eq. (3.4), i.e., primitive asset excess returns are linearly related to the excess return on the benchmark (market index) portfolio. This is not an innocuous assumption. A manager who invests in call options on the market will show spurious market-timing ability, since the value of a call option is a convex function of the return on the market. Such a manager will also show negative timing, as the following example, taken from Jagannathan and Korajczyk (1986), illustrates.

Consider a manager who buys the following one-period European call option on the total return market index (i.e., the index assumes reinvestment of dividends) at the beginning of each period. Let the value of the total return market index portfolio at time t be denoted by V_t. The call option has an exercise price, $K_t = V_t(1 + r_{f,t+1})$, and trades at price C_t. Let c_t and p_t denote the values of call and put options on the total return of the market index when the index value is 1.0 and the exercise price is $(1 + r_{f,t+1})$. The excess return on the call option, $R_{p,t+1}$, is given by

$$
\begin{aligned}
R_{p,t+1} &= \text{Max}\left(\frac{V_{t+1} - V_t(1 + r_{f,t+1})}{C_t}, 0\right) - (1 + r_{f,t+1}) \\
&= \frac{V_t}{C_t}\text{Max}\left(\frac{V_{t+1} - V_t(1 + r_{f,t+1})}{V_t}, 0\right) - (1 + r_{f,t+1}) \\
&= \frac{V_t}{C_t}\text{Max}(R_{m,t+1}, 0) - (1 + r_{f,t+1}) \\
&= -(1 + r_{f,t+1}) + \frac{V_t}{C_t}R_{m,t+1} - \frac{V_t}{C_t}\text{Min}(R_{m,t+1}, 0) \\
&= -(1 + r_{f,t+1}) + \frac{V_t}{C_t}R_{m,t+1} + \frac{V_t}{C_t}\text{Max}(-R_{m,t+1}, 0) \\
&\equiv a_p^S + \beta_{p,m}^U R_{m,t+1} + \beta_{p,m}^{U-D}\text{Max}[0, -R_{m,t+1}] + 0.
\end{aligned}
$$

Thus, $\beta_{p,m}^U = \beta_{p,m}^{U-D} = \frac{V_t}{C_t}$, and $a_p^S = -(1 + r_{f,t+1})$. The excess return on the manager's portfolio exactly fits Eq. (3.19), which measures the selection and timing skills of a manager. The value of selection skill will be the present value of $-(1 + r_{f,t+1})$ received one period from now (i.e., -1), and the value of timing skill equals the value of $\frac{V_t}{C_t}$ one-period put options on the total return market index with an index value of 1.0 and an exercise price of $(1 + r_{f,t+1})$, that is, $\frac{V_t}{C_t} \times p_t$. By Theorem 6 of Merton (1973a), this is also equal to the value of $\frac{1}{c_t}$ one-period put options on the total return market index when the index value is 1 and an exercise price of $(1 + r_{f,t+1})$, or $\frac{p_t}{c_t}$. Even though the manager is not doing any selection or timing, the managed portfolio returns exhibit positive measured timing skill and negative measured selection skill when evaluated using the Henriksson and Merton (1981) model. It can readily be verified, using the put–call parity theorem for European options, that $p_t = c_t$ and, therefore, the value of the portfolio's pseudo timing is 1. Hence, the excess value provided by the manger is 0, i.e., the value of timing and the value of selectivity exactly offset each other, and the manger does not add any value. While this one-for-one trade-off between timing and selection skills for a manager with no timing skills does not generalize to investing in options with different exercise prices, it is possible to bound the value of the spurious timing and selection skills created by a manager with no timing ability.

As Jagannathan and Korajczyk (1986) point out, the returns on certain asset classes have embedded call option-like features, and therefore a portfolio manager need not invest directly in options in order to exhibit spurious selection and timing skills. They find that a manager who holds a passive, equally weighted stock index portfolio shows significant negative selection skill and positive timing skill when the value weighted stock index is used as the market index portfolio.

A number of hedge-fund strategies appear to provide nonlinear payoff structures. This might be due to a number of factors: superior timing ability, direct holding of option positions, dynamic trading strategies that mimic option positions, or strategies that are equivalent to either buying or selling insurance. Mitchell and Pulvino (2001) show that a merger-arbitrage investment strategy looks very much like a short position in a put option on the market portfolio. Asness et al. (2001) and Connor et al. (2010) find that the preponderance of hedge-fund indices they study demonstrate higher down-market betas than up-market betas. Figure 3.1 illustrates this by plotting the monthly returns on the Credit Suisse Event-Driven Hedge-Fund Index against the monthly return on the S&P 500 Index over the period from January 1994 to August 2009. The piecewise-linear relation plotted in the figure is the fitted Henriksson–Merton timing regression (3.19). The estimated parameters are $a_p^S = 0.0087$ (10.44 % annualized, t-statistic = 5.35), $\beta_{p,m}^U = 0.08$ (t-statistic = 1.80), and $\beta_{p,m}^{U-D} = -0.26$ (t-statistic = −3.54). The patterns in the returns to event-driven strategies look similar to selling insurance on the S&P 500, with the premiums reflected in the measured security selection, a_p^S.

3.1.3.2 Frequent Trading and Pseudo Timing

Much of the market-timing literature implicitly assumes that the interval over which the observer measures returns corresponds to the portfolio-rebalancing period of the portfolio manager. That is, if we observe portfolio returns on a monthly basis, then the manager rebalances on a monthly basis and at the time we observe returns. In actuality, many active portfolio managers are likely to rebalance on a daily, or intra-daily, basis. Pfleiderer and Bhattacharya (1983) consider an example of

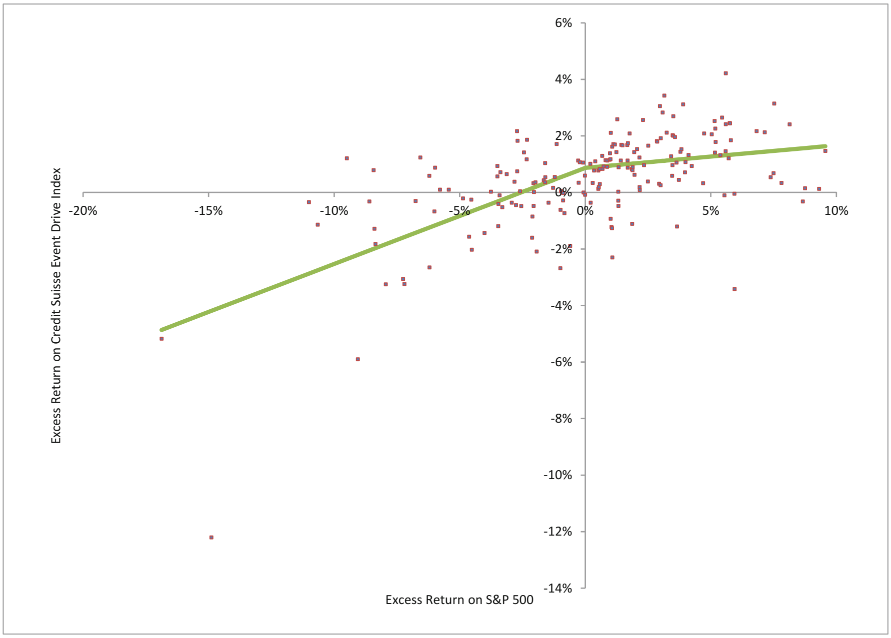


Fig. 3.1 Monthly excess returns on the Credit Suisse Event-Driven Hedge-Fund Index versus the excess returns of the S&P 500 Index. The piecewise-linear relation is the fitted Henriksson–Merton timing regression (Eq. 3.18). The regression parameter estimates are $\alpha^S = 0.0087$ (10.44 % annualized), $\hat{\beta}^U = 0.08$, and $\hat{\beta}^{U-D} = -0.26$

a manager who has no timing skill but rebalances the portfolio more frequently than the observation interval for portfolio returns. This manager is a "chartist" who bases the portfolio's beta on past market returns. The manager adjusts the portfolio positions each period, but portfolio returns are observed every second period. The return on the market from period t to $t + 2$ is

$$R_{m,t,t+2} = (1 + E[R_m] + \Delta_{m,t+1}) \times (1 + E[R_m] + \Delta_{m,t+2}) - 1. \tag{3.20}$$

If the chartist chooses exposure to the market to be a linear function of the lagged values of Δ, the chartist's two-period return is

$$\begin{aligned} r_{c,t,t+2} = {} & [1+r_f+(b_1+b_2\Delta_{m,t})(E[R_m]+\Delta_{m,t+1})] \\ & \times[1+r_f+(b_1+b_2\Delta_{m,t+1})(E[R_m]+\Delta_{m,t+2})]. \end{aligned} \tag{3.21}$$

This involves linear functions of $\Delta_{m,t+1}$ and $\Delta_{m,t+2}$ and a quadratic term in $\Delta_{m,t+1}$ which will yield a positive regression coefficient on the squared market return if b_2 is positive [see Pfleiderer and Bhattacharya (1983, Sect. 3)]. Since the chartist is creating measured timing ability without any true skill, the apparent timing ability is accompanied by negative measured security selection (α_p^S). Thus, portfolio rebalancing at a higher frequency than the observation interval used to evaluate performance causes the same type of difficulty that positions in derivatives or dynamic trading strategies designed limit losses through synthetic portfolio insurance. Hence, it will be difficult to identify true timing and selection skills based only on observations of the managed portfolio returns. One way to detect this type of pseudo timing is to study the relation of multi-period fund return, $r_{c,t,t+2}$, to the higher-frequency market returns. For the pseudo-timing chartist in this example, $r_{c,t,t+2}$ is related to $\Delta^2_{m,t+1}$ but not to $\Delta^2_{m,t+2}$ while a manager with true timing skill would have portfolio returns positively correlated with both $\Delta^2_{m,t+1}$ and $\Delta^2_{m,t+2}$.

Pfleiderer and Bhattacharya (1983) propose an approach that utilizes the fact that one often has access to higher-frequency returns on the market or benchmark portfolios even when the fund returns are observed infrequently. In the example above, the chartist's return $r_{c,t,t+2}$ will be correlated with $R_{m,t+1}$ but not with $R_{m,t+2}$. An alternative approach, proposed by Ferson et al. (2006),

3.1.4 A Contingent Claims Framework for Valuing the Skills of a Portfolio Manager

Glosten and Jagannathan (1994) show that the approach in Jagannathan and Korajczyk (1986) can be generalized to provide a consistent estimate of the value added by a portfolio manager due to true as well as pseudo selection and timing skills when taken together. In order to assess the value added by a portfolio manager when the portfolio return exhibits option-like features, Glosten and Jagannathan (1994), following Connor (1984), assume that the intertemporal marginal rate of substitution of consumption today for consumption tomorrow of the investor evaluating the abilities of the active portfolio manager is a time-invariant function of the returns on a few selected asset class portfolios. When the dynamic version of the Rubinstein (1976) CAPM holds, there will be only one asset class portfolio and it will be the return on the aggregate market portfolio.

The Glosten and Jagannathan (1994) approach involves regressing the excess return of the managed portfolio on the excess return on the market index portfolio and J one-period options on the market index portfolio, corresponding to J different exercise prices as given below:

$$r_{p,t} = \alpha + \beta r_{m,t} + \sum_{j=1}^{J} \gamma_j \mathrm{Max}(r_{m,t} - K_j, 0) + \varepsilon_{p,t},$$

where K_1 is set equal to 0, and the other $J - 1$ options are chosen judiciously so that $\alpha + \beta r_{m,t} + \sum_{j=1}^{J} \gamma_j \mathrm{Max}(r_{m,t} - K_j, 0)$ best approximates the return on the managed portfolio, $r_{p,t}$, for some choice of the parameters $\alpha, \beta, \gamma_j, j = 1, ..J$. They show that the average value of the manager's skill embodied in $r_{p,t}$ can be reasonably well approximated by $\frac{\alpha}{1+r_{f,t}} + \sum_{j=2}^{J} \beta_j C_j$, where C_j is the average value of the one-period option that pays $\text{Max}(r_{m,t} - K_j, 0)$ at time t by suitably choosing the number of options and their exercise prices. The valuation approach in Jagannathan and Korajczyk (1986) corresponds to $J = 1$ and $K_1 = r_{f,t}$.

With the advent of hedge funds, investors have access to portfolio managers who either directly invest in derivative securities or engage in trading behavior that create option-like features in their returns. As mentioned earlier, Mitchell and Pulvino (2001) show that the return on merger arbitrage, one particular hedge-fund strategy, has some of the characteristics of a written put option on the market portfolio. Fung and Hsieh (2001) show that the return on CTAs, another commonly used hedge-fund strategy, resembles the return on look-back options. They develop several benchmark returns that include returns on judiciously chosen options on several asset classes that are particularly suitable for assessing the performance of hedge-fund managers, and they are widely used in the academic literature as well as in practice. These methods build on the generalized Henriksson and Merton (1981) framework in Glosten and Jagannathan (1994). Ferson et al. (2006) provide an alternative way of addressing these issues.

3.1.5 Timing and Selection with Return Predictability

In our original formulation, deviations of market returns from their unconditional mean come from two sources: (a) deviations of market returns from their conditional mean (true shocks about which skilled managers may have forecasting ability), and (b) time variation in the conditional mean; $\Delta_{m,t} = \delta_{m,t} + \delta^*_{m,t}$ in (3.10). For simplicity of exposition, we have assumed that market portfolio returns are unpredictable from public information (i.e., $\delta^*_{m,t} = 0$). A large literature provides evidence for predictable time variation in the equity risk premium [e.g., Rozeff (1984), Keim and Stambaugh (1986), Campbell (1987), Campbell and Shiller (1988, 1989), Fama and French (1988, 1989), Breen et al. (1989)]. Cochrane (2011) observes that "Now it seems all price-dividend variation corresponds to discount-rate variation," although there is some debate about the predictability of market returns [e.g., Goyal and Welch (2003), Welch and Goyal (2008), Neuhierl and Schlusche (2011), and Cornell (2014)].

Ferson and Schadt (1996) show how to measure timing and selection when returns have a predictable component based on publicly available information. They start with the assumption that the conditional version of Eq. (3.2) and hence the conditional version of Eq. (3.4) hold, i.e.,

$$R_{i,t} = \beta_{i,m}(Z_{t-1})R_{m,t} + u_{i,t} \tag{3.22}$$

$$E(u_{i,t}|Z_{t-1}) = 0 \tag{3.23}$$

$$E(u_{i,t}R_{m,t}|Z_{t-1}) = 0, \tag{3.24}$$

where Z_{t-1} is a vector of instrumental variables that represent the information available at time $t-1$, $\beta_{i,m}(Z_{t-1})$ denotes the functional dependence of $\beta_{i,m}$ on Z_{t-1}, and $E(.|Z_{t-1})$ denotes the conditional expectations operator based on observing the vector of instrumental variables Z_{t-1}. Ferson and Schadt (1996) assume that the function $\beta_{i,m}(Z_{t-1})$ can be approximated well by $\beta_{i,m}(Z_{t-1}) = b_{0p} + B'_p z_{t-1}$, where $z_{t-1} = Z_{t-1} - E(Z_{t-1})$. With this additional assumption, Ferson and Schadt (1996) derive a conditional version of the Treynor and Mazuy (1966) model for detecting timing ability:

$$R_{p,t} = a_p^S + b_p R_{m,t} + B'_p z_{t-1} R_{m,t} + \gamma_p R_{m,t}^2 + \varepsilon_{p,t}. \tag{3.25}$$

B'_p captures the response of the manager's beta to the public information, γ captures the sensitivity of the manager's beta to the private market-timing signal, and a_p^S is a measure of the selection ability of the manager. They show that the following conditional version of the Henriksson–Merton model also holds:

$$\begin{aligned} R_{p,t} = {} & a_p^S + \beta_d R_{m,t} + B'_d z_{t-1} R_{m,t} + \gamma_c R_{m,t} I_{\{R_{m,t} - E(R_{m,t}|z_{t-1})>0\}} \\ & + \Delta' z_{t-1} R_{m,t} I_{\{R_{m,t} - E(R_{m,t}|z_{t-1})>0\}} + \varepsilon_{p,t}, \end{aligned} \tag{3.26}$$

where the function $I_{\{R_{m,t} - E(R_{m,t}|z_{t-1})>0\}}$ takes the value of 1 when $R_{m,t} - E(R_{m,t}|z_{t-1}) > 0$ and 0 otherwise.

Using monthly return data on 67 mutual funds during 1968–1990, Ferson and Schadt (1996) find that the risk exposure of mutual funds changes in response to publicly available information on the stock index dividend yield, short-term interest rate, slope of the treasury yield curve, and corporate-bond yield spread. Unlike the unconditional Jensen's alpha (selection measure), which is negative on average across funds, the conditional selection measure is on average zero. When the conditional models in Eqs. (3.26) and (3.25) are used, the perverse market timing exhibited by US mutual funds goes away, highlighting the need for controlling for predictable components in stock returns. However, the data pose an interesting puzzle since managers seem to pick market exposures that are positively correlated with $\delta_{m,t}$ but negatively correlated with $\delta^*_{m,t}$. Ferson and Warther (1996) show evidence indicating that the anomalous negative correlation between fund betas and $\delta^*_{m,t}$ is caused by flows of funds into mutual funds prior to high market returns. Delay in allocating those funds from cash to other assets causes a drop in beta prior to high return periods. Christopherson et al. (1998) find that alphas do not differ between conditional and unconditional performance measures for pension fund portfolios. This is consistent with the fund flows argument for perverse timing for mutual funds if pension funds are less subject to fund flows that are correlated with $\delta^*_{m,t}$. Ferson and Qian (2004) expand the time period and cross-sectional sample of funds studies and find results that are broadly consistent with the earlier conditional performance evaluation literature.

3.2 Holdings-Based Performance Measurement

We have focussed on returns-based performance evaluation of market timing, in the spirit of Treynor and Mazuy (1966). When the fund manager's portfolio holdings are observed, the investor can use that additional information in measuring the timing and selection abilities of the manager with more precision. There is a vast literature on holdings-based performance evaluation going back, at least, to Fama (1972). Holdings-based performance evaluation is quite common in practice when the investor is in a position to see the portfolio positions on a frequent basis. A common practice in industry is to attribute the performance difference between the managed portfolio and the benchmark into two components: that due to "allocation" and that due to "selection". As discussed in Sharpe (1991), *allocation* takes the weights assigned by the manager to the different sectors and compare them with the weights for those sectors in the benchmark, and computes the effect of those deviations from benchmark allocation weights. The residual is classified as *selection*. Daniel et al. (1997) build on these practices to decompose the return on an actively managed portfolio into three components: characteristics-based selection, characteristics-based timing, and average characteristics-based style. We cannot do justice to that literature here but briefly touch on it.

Following Kacperczyk et al. (2014), define timing as follows:

$$Timing_{j,t} = \sum_{i=1}^{N^j} \left(w_{i,t}^j - w_{i,t}^m \right) \beta_{i,t} R_{m,t+1} = (\beta_{p,t}^j - 1) R_{m,t+1} \tag{3.27}$$

where $Timing_{j,t}$ is the timing skill of manager j at time t, $w_{i,t}^j$ is the weight of security i at time t in manager j's portfolio, $w_{i,t}^m$ is the corresponding security's weight in the market portfolio, N_t^j is the number of securities in manager j's portfolio at time t, and $\beta_{i,t}$ is the covariance of security i's excess return with the excess return on the market portfolio divided by the variance of the market portfolio's excess return based on information available at time t. Note that $\left(w_{i,t}^j - w_{i,t}^m \right) \beta_{i,t}$ is multiplied by $R_{m,t+1}$, the excess return on the market portfolio at time $t + 1$. We would expect that a manager with timing ability would construct the portfolio such that there is positive correlation between $\beta_{p,t}^j$ and $R_{m,t+1}$. In a similar way, define the selection skill as

$$\text{Selection}_{j,t} = \sum_{i=1}^{N^j} \left(w_{i,t}^j - w_{i,t}^m \right) (R_{i,t+1} - \beta_{i,t} R_{m,t+1}) = \varepsilon_{p,t}^j. \tag{3.28}$$

Notice that observing the portfolio holdings of the fund manager facilitates measuring the systematic risk exposure of the manager's portfolio at any given point in time, t, more precisely. When the weights assigned to the securities in the manager's portfolio changes over time, the managed fund's beta will also vary over time even when the betas of individual securities remain constant. Hence, the holdings information helps assess the manager's abilities better.

Kacperczyk et al. (2014) estimate the selection and timing skills of a manager using the following time-series regressions:

$$\text{Timing}_{j,t} = a_0 + a_1 \text{Recession}_t + a_2 X_t^j + u_t^j \tag{3.29}$$

$$\text{Selection}_{j,t} = b_0 + b_1 \text{Recession}_t + b_2 X_t^j + v_t^j, \tag{3.30}$$

where Recession_t^j is an indicator variable equal to one if the economy in month t is in a recession as defined by NBER and zero otherwise. X_t^j is a set of fund-specific control variables, including age, size, expense ratio, turnover, percentage flow of new money, load fees, other fees, and other fund style characteristics. The use of a recession dummy variable is based on the evidence that the equity premium is countercyclical. The authors find that a subset of managers do have superior skills. The same managers exhibit both superior timing and selection skills. Superior performance due to timing is more likely during recessions while selection is dominant during other periods.

3.3 Summary

Most individual and institutional investors rely on professional money managers. While delegation provides gains through specialization, it also imposes invisible agency costs: an investor has to evaluate managers as well as select and monitor the ones with superior skills. Return-based and portfolio holdings-based performance measures complement each other in identifying portfolio managers with superior abilities. While we briefly cover holdings-based performance measures, our main focus is on return-based performance measures. Treynor (1965) and Treynor and Mazuy (1966) are the earliest examples of return-based performance measures assuming constant and variable exposures to market risk, respectively.

Measuring performance requires a conceptual framework. The literature has evolved by examining whether a representative investor would benefit from access to an actively managed fund. In order to answer the question, we need to know the objectives of the representative investor and which portfolio the investor will hold in the absence of access to the active manager. The CAPM provides a natural starting point. According to the CAPM, investors care only about the mean and the variance of the return on their wealth portfolio and, furthermore, the representative investor will hold the market portfolio of all assets. Treynor (1965) and Jensen (1968) develop security-selection ability measures that are valid from the perspective of such a representative investor. These measures assess value at the margin for a small incremental investment in the active fund from the perspective of the representative investor holding the market portfolio. The value at the margin is not sufficient for deciding how to allocate funds across the market and actively managed funds, since allocating nontrivial amounts in active funds will lead to the investor's portfolio deviating significantly from the market portfolio, and the marginal valuations will change. Treynor and Black (1973) solve for optimal portfolio choice when funds or individual assets provide abnormal returns, creating a framework for asset allocation that provides the conceptual foundation for many modern-day quantitative investment strategies.

The performance measures of Treynor (1965) and Jensen (1968) and the asset-allocation framework of Treynor and Black (1973) assume, in addition to the assumptions leading to the CAPM, that security returns are linearly related to the return on the market index portfolio; i.e., the up and down market betas of a security or managed portfolio are the same. Treynor and Mazuy (1966) make the important observation that the return on the portfolio of a fund manager who successfully forecasts market returns and adjusts market exposure will resemble the return on a call option and will be nonlinearly related to the return on the market with higher beta in up markets. Dybvig and Ingersoll (1982) show that while the CAPM may provide a reasonable framework for valuing major asset classes, and securities whose returns are linearly related to the market return, the CAPM will in general assign the wrong value to payoffs with option-like features. Merton (1981) addresses this issue by developing a framework for assessing the value, at the margin, of a successful market-timing fund manager. Pfleiderer and Bhattacharya (1983) and Admati et al. (1986) provide additional insights by modeling the behavior of a fund manager with access to informative signals. Their analysis shows the rather restrictive nature of the assumptions that are needed to support the dichotomous classification of the abilities of a fund manager into timing and selection.

Jagannathan and Korajczyk (1986) and Glosten and Jagannathan (1994) show that, while the timing and selection skills of an active fund manager cannot be disentangled, classifying the skill into two types—selection and market timing (or asset allocation)—facilitates assessing the correct value, at the margin, created by the manager; the combined value of timing and asset selection is the crucial variable of interest, and that can be assessed with reasonable precision. While zero-value strategies that involve taking positions in securities with option-like features or mimicking timing skill through frequent rebalancing can show superior performance in one dimension (timing or selection), that performance will come at the expense of poor performance in the other dimension. Properly disentangling selection and performance even when they are spurious allows the investor to measure the net value added by a fund manager more precisely.

All these performance-evaluation methods assume that security returns have little predictability based on publicly available information. While the literature documenting predictable patterns in stock returns based on publicly available information has become rather large, the practical relevance of the findings is the subject of ongoing debate. In a world where risk premiums can be forecasted with publicly available information, there is a need for performance measures that distinguish between true timing ability and timing through the use of public information. Ferson and Schadt (1996) derive measures that allow us to isolate true skill.

References

Admati, A. R., Bhattacharya, S., Pfleiderer, P., & Ross, S. A. (1986). On timing and selectivity. *Journal of Finance, 41*, 715–730.

Asness, C., Krail, R., & Liew, J. (2001). Do hedge funds hedge? *Journal of Portfolio Management, 28*, 6–19.

Board of Governors of the Federal Reserve System. (2014). *Financial accounts of the United States: Flow of funds, balance sheets, and integrated macroeconomic accounts, first quarter 2014*. Washington: Board of Governors of the Federal Reserve System, http://www.federalreserve.gov/releases/z1/Current/z1.pdf.

Bodie, Z., Kane, A., & Marcus, A. J. (2011). *Investments* (9th ed.). New York: McGraw Hill/Irwin.

Breen, W., Glosten, L. R., Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *Journal of Finance, 44*, 1177–1189.

Busse, J. A. (1999). Volatility timing in mutual funds: Evidence from daily returns. *Review of Financial Studies, 12*, 1009–1041.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics, 18*, 373–399.

Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *Journal of Finance, 43*, 661–676.

Campbell, J. Y., & Shiller, R. J. (1989). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*, 195–228.

Chen, Y., Ferson, W., & Peters, H. (2010). Measuring the timing ability and performance of bond mutual funds. *Journal of Financial Economics, 98*, 72–89.

Chen, Y., & Liang, B. (2007). Do market timing hedge funds time the market? *Journal of Financial and Quantitative Analysis, 42*, 827–856.

Christopherson, J. A., Ferson, W., & Glassman, D. A. (1998). Conditioning manager alpha on economic information: Another look at the persistence of performance. *Review of Financial Studies, 11*, 111–142.

Cochrane, J. H. (2011). Discount rates. *Journal of Finance, 66*, 1047–1108.

Connor, G. (1984). A unified beta pricing theory. *Journal of Economic Theory, 34*, 13–31.

Connor, G., Goldberg, L. R., & Korajczyk, R. A. (2010). *Portfolio risk analysis*. Princeton: Princeton University Press.

Cornell, B. (2014). Dividend-price ratios and stock returns: International evidence. *Journal of Portfolio Management, 40*, 122–127.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance, 52*, 1035–1058.

Dybvig, P. H., & Ingersoll, J. E., Jr. (1982). Mean-variance theory in complete markets. *Journal of Business, 55*, 233–251.

Dybvig, P. H., & Ross, S. A. (1985). Differential information and performance measurement using a security market line. *Journal of Finance, 40*, 383–399.

Elton, E. J., Gruber, M. J., & Busse, J. A. (2004). Are investors rational? Choices among index funds. *Journal of Finance, 59*, 261–288.

Fama, E. F. (1972). Components of investment performance. *Journal of Finance, 27*, 551–68.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–25.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*, 23–49.

Ferson, W., Henry, T. R., & Kisgen, D. J. (2006). Evaluating government bond fund performance with stochastic discount factors. *Review of Financial Studies, 19*, 423–455.

Ferson, W., & Mo, H. (2016). Performance measurement with market and volatility timing and selectivity. *Journal of Financial Economics, 121*, 93–110.

Ferson, W. E., & Qian, M. (2004). *Conditional performance evaluation, revisited*. Charlottesville, VA: CFA Institute Research Foundation.

Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic times. *Journal of Finance, 51*, 425–461.

Ferson, W. E., & Warther, V. A. (1996). Evaluating fund performance in a dynamic market. *Financial Analysts Journal, 52*, 20–28.

French, C., & Ko, D. (2006). *How hedge funds beat the market*, Working paper, http://ssrn.com/abstract=927235.

Fung, W., & Hsieh, D. A. (2001). The risk in hedge fund strategies: Theory and evidence from trend followers. *Review of Financial studies, 14*, 313–341.

Glosten, L. R., & Jagannathan, R. (1994). A contingent claim approach to performance evaluation. *Journal of Empirical Finance, 1*, 133–160.

Goyal, A., & Welch, I. (2003). Predicting the equity premium with dividend ratios. *Management Science, 49*, 639–654.

Grant, D. (1977). Portfolio performance and the “cost” of timing decisions. *Journal of Finance, 32*, 837–846.

Hallerbach, W. G. (2014). On the expected performance of market timing strategies. *Journal of Portfolio Management, 40*, 42–51.

Henriksson, R. D. (1984). Market timing and mutual fund performance: An empirical investigation. *Journal of Business, 57*, 73–96.

Henriksson, R. D. & Merton, R. C. (1981). On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. *Journal of Business, 54*, 513–533.

Investment Company Institute. (2014). *Investment company fact book* (54th ed.). Washington: Investment Company Institute, http://icifactbook.org/pdf/2014_factbook.pdf.

Jagannathan, R., & Korajczyk, R. A. (1986). Assessing the market timing performance of managed portfolios. *Journal of Business, 59*, 217–235.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance, 23*, 389–416.

Jensen, M. C. (1972). Optimal utilization of market forecasts and the evaluation of investment performance. In G. P. Szegö & K. Shell (Eds.), *Mathematical methods in investment and finance*. Amsterdam: North-Holland.

Kacperczyk, M., Van Nieuwerburgh, S., & Veldkamp, L. (2014). Time-varying fund manager skill. *Journal of Finance, 69*, 1455–1484.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics, 17*, 357–390.

Kon, S. J. (1983). The market timing performance of mutual fund managers. *Journal of Business, 56*, 323–347.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13–37.

Merton, R. C. (1973a). Theory of rational option pricing. *Bell Journal of Economics and Management Science, 4*, 141–183.

Merton, R. C. (1973b). An intertemporal capital asset pricing model. *Econometrica, 41*, 867–887.
Merton, R. C. (1981). On market timing and investment performance I: An equilibrium theory of value for market forecasts. *Journal of Business, 54*, 363–406.
Mitchell, M., & Pulvino, T. (2001). Characteristics of risk and return in risk arbitrage. *Journal of Finance, 56*, 2135–2175.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*, 768–783.
Neuhierl, A., & Schlusche, B. (2011). Data snooping and market-timing rule performance. *Journal of Financial Econometrics, 9*, 550–587.
Pfleiderer, P., & Bhattacharya, S. (1983). *A note on performance evaluation*, Technical report #714, Stanford University.
Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341–360.
Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *Journal Portfolio Management, 11*, 68–75.
Rubinstein, M. (1976). The valuation of uncertain income streams and the pricing of options. *Bell Journal of Economics, 7*, 407–425.
Sharpe, W. F. (1964). A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Sharpe, W. F. (1991). *The AT&T pension fund*, Stanford Business School Case SF-243.
Treynor, J. L. (1962). *Toward a theory of market value of risky assets*. Unpublished manuscript.
Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.
Treynor, J. L. (1999). Toward a theory of market value of risky assets. In: R. A. Korajczyk, (Ed.), *Asset pricing and portfolio performance*. London: Risk Books.
Treynor, J. L., & Black, F. (1973). How to use security analysis to improve portfolio selection. *Journal of Business, 46*, 66–86.
Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131–136.
Viceira, L. M. (2009). Life-cycle funds. In: A. Lusardi, (Ed.), *Overcoming the saving slump: How to increase the effectiveness of financial education and saving programs*. Chicago: University of Chicago Press.
Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction.' *Review of Financial Studies, 21*, 1455–1508.

Chapter 4
Returns, Risk, Portfolio Selection, and Evaluation

Phoebus J. Dhrymes and John B. Guerard

We present additional evidence on the risk and return of stocks in the USA and globally in the 1997–2009 period. We use a stock selection model incorporating fundamental data, momentum, and analysts' expectations and create portfolios using fundamental and statistically based risk models. We find additional evidence to support the use of multifactor models for portfolio construction and risk control. We created portfolios for the January 1997 to December 2009 period. We report three conclusions: (1) a stock selection model incorporating reported fundamental data, such as earnings, book value, cash flow, and sales, and analysts' earnings forecasts and revisions and momentum can identify mispriced securities; (2) statistically based risk models produce a more effective return-to-risk portfolio measures than fundamentally based risk models; and (3) the global portfolio returns of the multifactor risk-controlled portfolio returns dominate USA-only portfolios.

4.1 Introduction and Summary

How do we conceptualize the operation of markets for risky assets? To do so we need to take into account of: how the price of individual assets is determined, how risks and returns are balanced and how groups of such assets (portfolios) are put

P.J. Dhrymes (✉)
Department of Economics, Columbia University, New York, NY, USA

J.B. Guerard
McKinley Capital Management, LLC, Anchorage, AK, USA
e-mail: jguerard@McKinleyCapital.com

© Springer International Publishing Switzerland 2017
J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_4

together, i.e., the principles by which they are constructed, how they balance returns and risks, and once they are brought into being how their performance may be evaluated.

Evidently, this task begins with the evaluation of what one might expect the returns of a risky asset to be. Expected returns on assets are not completely explained by using only historical means (and standard deviations). One can estimate models of expected return using earnings expectations data, price momentum variables and other relevant financial data. In this analysis, we construct and estimate a stock selection model using US stocks and global stocks incorporating reported financial data earnings, expectations data, and price momentum for the period from January 1985 to December 2009. Despite the recent volatility of the momentum factor, momentum is still statistically associated with security returns and can be used with other factors to rank-order stocks for purchase. A composite value of momentum, value, and growth factors is estimated for US equities universe to identify potentially mispriced stocks. In addition, we consider the regression-weighting of factors, enhanced information coefficients relative to equally weighted factors. Thus, momentum and analysts' forecast variables dominate the regression-based composite model of expected returns. We created portfolios for the January 1997 to December 2009 period. We report three conclusions: We report three conclusions: (1) a stock selection model incorporating reported fundamental data, such as earnings, book value, cash flow, sales, analysts' earnings forecasts, revisions and momentum can identify mispriced securities; (2) statistically based risk models produce a more effective return-to-risk portfolio measures than fundamentally based risk models; and (3) the global portfolio returns of the multifactor risk-controlled portfolio returns dominate USA-only portfolios. In this study, we review recent expected returns modeling literature, trace the development of enhanced multifactor mean-variance portfolio construction models, and show domestic and global portfolio management and analysis.

4.2 Expected Returns Modeling and Stock Selection Models: Recent Evidence

There are many approaches to security valuation and the creation of expected returns. The first approaches to security analysis and stock selection involved the use of valuation techniques using reported earnings and other financial data. Graham and Dodd (1934) recommended that stocks be purchased on the basis of the price-earnings (*P/E*) ratio. They suggested that no stock should be purchased if its price-earnings ratio exceeded 1.5 times the *P/E* multiple of the market. Graham and Dodd established the *P/E* criteria, which was discussed in Williams (1938), the monograph that influenced Harry Markowitz and his thinking on portfolio construction. It is interesting that Graham and Dodd put forth the low *P/E* model at the height of the Great Depression. Basu (1977) reported evidence supporting the low *P*/*E* model. Dremen (1979, 1998) presented practitioner evidence of low *P*/*E* effectiveness. Academics often prefer to test the low *P*/*E* approach by testing its reciprocal, the "high *E/P*" approach. The high *E/P* approach specifically addresses the issue of negative earnings per share, which can confuse the low *P*/*E* test. Hawawini and Keim (1995) found statistical support for the high *EP* variable of NYSE and AMEX stocks from April 1962 to December 1989. At a minimum, Graham and Dodd also advocated the calculation of a security's net current asset value, NCAV, defined as its current assets less all liabilities. A security should be purchased if its net current value exceeded its current stock price. The price-to-book (*PB*) ratio should be calculated, but not used as a measure for stock selection.

There is extensive literature on the impact of individual value ratios on the cross section of stock returns. We go beyond using just one or two of the standard value ratios (*EP* and *BP*) to include the cash-price ratio (*CP*) and/or the sales-price ratio (*SP*). Research on the combination of value ratios to predict stock returns that include at least *CP* and/or *SP* include Chan, Hamao, and Lakonishok (1991), Bloch, Guerard, Markowitz, Todd, and Xu (1993), Lakonishok, Shleifer, and Vishny (1994), and Guerard, Gultekin, and Stone (1997), Haugen and Baker (2010), and Stone and Guerard (2010). We review these seven papers in some detail and provide the following comments.

Chan et al. (1991) used seemingly unrelated regression (SUR) to model CAPM monthly excess returns as functions of CAPM excess returns of the value-weighted or equal-weighted market index return; *EP*, *BP*, *CP*; size as measured by the natural logarithm of market capitalization (LS).[1] Betas were simultaneously estimated and cross-sectional correlations of residuals were addressed. When fractile portfolios were constructed by sorting on the *EP* ratio, the highest *EP* quintile portfolio outperformed the lowest *EP* quintile portfolio, and the *EP* effect was not statistically significant. The highest *BP* stocks outperformed the lowest *BP* stocks. The portfolios composed (sorted) of the highest *BP* and *CP* outperformed the portfolios composed of the lowest *BP* and *CP* stocks. In the authors' multiple regressions, the size and book-to-market variables were positive and statistically significant. The *EP* coefficient was negative and statistically significant at the 10 % level. Thus, no support was found for the Graham and Dodd low *PE* approach. In the monthly univariate SUR analysis, with each month variable being deflated by an annual (June) cross-sectional mean, Chan et al. (1991) found that the *EP* coefficient was negative (but not statistically significant), the size coefficient was negative (but not statistically significant), the book to market coefficient was positive and statistically significant, and the cash flow coefficient was positive and statistically significant. In their multiple regressions, Chan et al. (1991) report *BP* and *CP* variables were positive and statistically significant but *EP* was not significant. Applying an adaptation of the Fama and MacBeth (1973) time series of portfolio cross sections to the Japanese market produced negative and statistically significant

[1]Chan et al. (1991) define cash as the sum of earnings and depreciation without explicit correction for other noncash revenue or expenses.

coefficients on *EP* and size but positive and statistically significant coefficients for the *BP* and *CP* variables. Chan et al. (1991, p. 1760) summarized their findings: "The performance of the book to market ratio is especially noteworthy; this variable is the most important of the four variables investigated."

In 1991, Markowitz headed the Daiwa Securities Trust Global Portfolio Research Department (GPRD). The Markowitz team estimated stock selection models using Graham et al. (1934) fundamental valuation variables, earnings, book value, cash flow and sales, relative variables, defined as the ratio of the absolute fundamental variable ratios divided by the 60-month averages of the fundamental variables. Bloch et al. (1993) reported a set of some 200 simulations of US and Japanese equity models. Bloch et al. (1993) found that Markowitz (1987) mean-variance efficient portfolios using the lower *EP* values in Japan underperformed the universe benchmark, whereas *BP*, *CP*, and *SP* (sales-to-price, or sales yield) variables outperformed the universe benchmark. For the US optimized portfolios using *BP*, *CP*, *SP*, and *EP* variables, the portfolios outperformed the U.S. S&P 500 index, giving support to the Graham and Dodd concept of the low price-to-earnings variable using the relative rankings of value-focused fundamental ratios to select stocks.[2] Bloch et al. (1993) used relative ratios as well as current ratio values. Not only might an investor want to purchase a low *P*/*E* stock, but one might wish to purchase a low *P*/*E* stock when the *P*/*E* is at a relatively low value compared to its historical value, in this case a low relative to its average over the last 5 years.

Let TR denote 3-month-ahead stock return for a stock. Bloch et al. (1993) estimate the following regression equation to assess empirically the relative explanatory power of each of the eight value ratios in the equation:

$$\begin{aligned}\mathrm{TR} = w_0 + w_1 EP + w_2 BP + w_3 CP + w_4 SP + w_5 REP \\ + w_6 RBP + w_7 RCP + w_8 RSP + e_t. \end{aligned} \tag{4.1}$$

Given concerns about both outlier distortion and multicollinearity, Bloch et al. (1993) tested the relative explanatory and predictive merits of alternative regression estimation procedures: OLS, robust using the Beaton and Tukey (1974) bi-square criterion to mitigate the impact of outliers, latent root to address the issue of multicollinearity [see Gunst, Webster, and Mason (1976)], and weighted latent root, denoted WLRR, a combination of robust and latent root. Bloch et al. (1993) used the estimated regression coefficients to construct a rolling horizon return forecast. The predicted returns and predictions of risk parameters were used as input to a mean-variance optimizer [see Markowitz (1987)] to create mean-variance efficient

[2] One finds the Price/Earnings, Price/Book, Price/Sales listed among the accounting anomalies in Levy (1999), p. 434. Levy also discusses the dividend yield as a (positive) stock anomaly. Malkiel (1996) cites evidence in support of buying low *P*/*E*, low *P*/*B*, and high *D*/*P* (dividend yield) stocks for outperformance, provided the low *P*/*E* stocks have modest growth prospects (pp. 204–210). Malkiel speaks of a "double bonus"; that is, if growth occurs, earnings increase and the price-to-earnings multiple may increase, further driving up the price. Of course, should growth fail to occur, both earnings and the *P*/*E* multiple may fall.

portfolios in both Japan (first section, nonfinancial Tokyo Stock Exchange common stocks, January 1975 to December 1990) and the USA (the 1000 largest market-capitalized common stocks, November 1975 to December 1990).[3]

Bloch et al. (1993) reported several results. First, for both Japanese and the US financial markets, they compared OLS and WLLR techniques, inputting expected returns forecasts produced by each method into a mean-variance optimizer. The WLRR-constructed composite model portfolio produced higher Sharpe Ratios and geometric means than the OLS-constructed composite variable portfolio in both Japan and the USA, indicating that controlling for both outliers and multicollinearity is important in using regression-estimated composite forecasts. Second, Bloch et al. quantified survivor bias and found it was not statistically significant in Japan and the USA for the period tested. Third, they investigated period-to-period portfolio revision and found that tighter turnover and rebalancing triggers led to higher portfolio returns for value-based strategies. Finally, Markowitz and Xu (1994) developed a test for data mining. In addition to testing the hypothesis of data mining, the test can be used to estimate and assess the expected difference between the best test model and the average of simulated policies.

In a thorough assessment of value versus growth in Japan and the USA, Lakonishok et al. (1994) examined the intersection of Compustat and CRSP databases for annual portfolios for NYSE and AMEX common stocks, April 1963 to April 1990. Their value measures were three current value ratios: *EP*, *BP*, and *CP*. Their growth measure was the 5-year average annual sales growth (GS). They performed three types of tests: a univariate ranking into annual decile portfolios for each of the four variables, bivariate rankings on *CP* (value) and GS (growth, glamor), and finally a multivariate regression adaptation of the Fama and MacBeth (1973) time series pooling of cross sectional regressions. Lakonishok et al. (1994) used the Fama-MacBeth methodology to construct portfolios and pool (average over time) a time series of twenty-two 1-year cross-sectional univariate regressions for each of the 22 years in their study period. The univariate regression coefficient for SG was significantly negative. The *EP*, *BP*, and *CP* coefficients were all significantly positive. When Lakonishok, Shleifer, and Vishny performed a multivariate regression using all four variables, they found significantly positive coefficients for *BP* and *EP* (but not *CP*) and significantly negative coefficients for SG. Overall, Lakonishok et al. (1994) concluded that buying out-of-favor value stocks outperformed growth (glamor) over the April 1968 to April 1990 period, that future growth was difficult to predict from past growth alone and that the actual

[3]The use of nonfinancial stocks led to a customized index for the Markowitz Global Portfolio Research Group (GPRD) analysis. The Chan et al. and an initial Guerard presentation occurred in September 1991 at the Berkeley Program in Finance, Santa Barbara, on Fundamental Analysis. Bill Ziemba presented a very interesting study comparing US and Japanese fundamental strategies at the same Berkeley Program meeting. Markowitz refers to this meeting in his Nobel Prize lecture (1991). Ziemba and Schwartz (1993) used capitalization-weighted regressions. The Chan et al., Guerard, and Ziemba studies found statistical significance with expectation and reported fundamental data.

future growth of the glamor stocks was much lower than past growth relative to the growth of value stocks, and that the value strategies ex post were not significantly riskier than growth (glamor) strategies.

Guerard et al. (1997) studied the intersection of Compustat, CRSP and I/B/E/S databases. This study built on the fundamental forecasting work in Bloch et al. (1993) in two ways: (1) adding to the Bloch et al. eight-variable regression equation a growth measure; and then (2) adding three measures of analysts' forecasts and forecast revisions from the I/B/E/S database, namely consensus analysts' forecasts, forecast revisions, and the direction (net up or down) of the forecast revisions. We will use the GSG (1997) consensus I/B/E/S variable priority growth variable, denoted PRGR in the original analysis. In quarterly weighted latent root regressions, the growth variable averaged a relative weight of 33 % whereas the average relative weighting of the 8 value variables averaged almost 67 %. In a 9-factor regression model, instead of having an average coefficient of 0.111, the PRGR variable had an average coefficient of 0.33, the largest weight in the model. The GGS result complements that of Lakonishok et al. (1994) in showing that rank-ordered portfolio returns have both a significant value and growth components.

Adding I/B/E/S variables to the eight value ratios produced more than 2.5 % of additional annualized return. The finding of significant predictive performance value for the three I/B/E/S variables indicates that analyst forecast information has value beyond purely statistical extrapolation of past value and growth measures. Possible reasons for the additional performance benefit could be that analysts' forecasts and forecast revisions reflect information in other return-pertinent variables, or discontinuities from past data, or serve as a quality screen on otherwise out-of-favor stocks. The quality screen idea would confirm Graham and Dodd's argument that value ratios should be used in the context of the many qualitative and quantitative factors that they argue are essential to informed investing. In terms of relative predictive value, Guerard et al. (1997) found the *EP*, *CP*, *SP*, and *RSP* variables to be more important than the *BP* variable. To test the risk-corrected performance value of the forecasts, Guerard et al. (1997) formed quarterly portfolios with risk being modeled via a 4-factor APT-based model (created using 5 years of past monthly data). The portfolios' quarterly returns averaged 6.18 % before correcting for risks and transaction costs with excess returns of 3.6 % after correcting for risk and 2.6 % quarterly after subtracting 100 basis points to reflect an estimate of two-way transactions costs.

Haugen and Baker (2010) extended their 1996 study in a recent volume to honor Harry Markowitz (Guerard, 2010). Haugen and Baker estimate their model using weighted least squares. In a given month we will simultaneously estimate the payoffs to a variety of firm and stock characteristics using a weighted least squares multiple regression procedure of the following form:

$$r_{j,t} = \sum_{i=1}^{n} P_{i,t} F_{i,j,t-1} + \mu_{j,t} \tag{4.2}$$

where:

$r_{j,t}$ = the total rate of return to stock j in month t.

$P_{i,t}$ = estimated weighted least squares regression coefficient (payoff) for factor i in month t.

$F_{i,j,t-1}$ = normalized value for factor i for stock j at the end of month $t-1$.

n = the number of factors in the expected return factor model.

$\mu_{j,t}$ = component of total monthly return for stock j in month t unexplained by the set of factors.

Haugen and Baker (2008) estimated their equation (4.2) in each month in the period 1963 through 2007.[4] In the manner of Fama and MacBeth (1973), they then compute the average values for the monthly regression coefficients (payoffs) across the entire period. Dividing the mean payoffs by their standard errors we obtain t-statistics. The most significant factors are computed as follows:

- Residual Return is last month's residual stock return unexplained by the market.
- Cash Flow-to-Price is the 12-month trailing cash flow-per-share divided by the current price.
- Earnings-to-Price is the 12-month trailing earnings-per-share divided by the current price.
- Return on Assets is the 12-month trailing total income divided by the most recently reported total assets.
- Residual Risk is the trailing variance of residual stock return unexplained by market return.
- 12-month Return is the total return for the stock over the trailing 12 months.
- Return on Equity is the 12-month trailing earnings-per-share divided by the most recently reported book equity.
- Volatility is the 24-month trailing volatility of total stock return.
- Book-to-Price is the most recently reported book value of equity divided by the current market price.
- Profit Margin is 12-month trailing earnings before interest divided by 12-month trailing sales.
- 3-month Return is the total return for the stock over the trailing 3 months.
- Sales-to-Price is 12-month trailing sales-per-share divided by the market price.

Haugen and Baker noted that the t-scores are large as compared to those obtained by Fama and MacBeth even though the length of the time periods covered by the studies is comparable.[5] Last month's residual return and the return over the preceding 3 months have negative predictive power relative to next month's total return. This may be induced by the fact that the market tends to overreact to most

[4] Fifty-seven factors are used in the model. See Haugen and Baker (1996) for definitions.

[5] In Fama and French (2008) p. 1668, ad hoc cross-section regressions are used in an attempt to explain the cross-sectional structure of stock returns. They report t-statistics as large as −8.59, but no attempt is made to investigate the out-of-sample predictive power of their regressions. Fama and French further research book-to-price and momentum anomalies in their 1995 and 2008 studies.

information. The overreaction sets up a tendency for the market to reverse itself upon the receipt of the next piece of related information.

The four measures of cheapness: cash-to-price, earnings-to-price, book-to-price, and sales-to-price, all have significant positive payoffs. Haugen and Baker (2010) find statistically significant results for the four fundamental factors as did the previously studies we reviewed. Measures of cheapness have been frequently found in the past[6] to be associated with relatively high stock returns, so it is not surprising that four measures of cheapness appear here as significant determinants of structure in the cross-section. Haugen and Baker (2010) dismiss the problem of multicollinearity.[7] Haugen and Baker present optimization analysis to support their stock selection modeling, and portfolio trading is controlled through a penalty function. When available, the optimizations are based on the largest 1000 stocks in the database. Estimates of portfolio volatility are based on the full covariance matrix of returns to the 1000 stocks in the previous 24 months. Two years of monthly return data, from 1963 through 1964, is used to construct the initial portfolios. Estimates of expected returns to the 1000 stocks are based on the factor model discussed above. The following constraints are applied to portfolio weights for each quarterly optimization:

1. The maximum weight in a portfolio that can be assigned to a single stock is limited to 5 %. The minimum is 0 % (Short selling is not permitted).
2. The maximum invested in any one stock in the portfolio is three times the market capitalization weight or 0.25 %, whichever is greater, subject to the 5 % limit.
3. The portfolio industry weight is restricted to be within 3 % of the market capitalization weight of that industry. (Based on the two-digit SIC code.)
4. Turnover in the portfolio is penalized through a linear cost applied to the trading of each stock. As a simplification, all stocks are subject to the same linear turnover cost although in practice portfolio managers use differential trading costs in their optimizations.

These constraints are designed to merely keep the portfolios diversified. Reasonable changes in the constraints do not materially affect the results. The portfolios are reoptimized quarterly.[8]

Trading costs are not reflected in the Haugen and Baker (2010) optimization analysis; however, the Haugen and Baker (2010) portfolios out-performed the benchmark by almost 5 % with average annual turnover of 80 % during the 1965–

[6]See, for example, Fama and French (1992).

[7]Haugen and Baker (2010) address the argument that these measures of cheapness in the regressions would make the methodology prone to multicolinearity. Significant problems associated with multicollinearity should result in instability in the estimated regression coefficients from month to month. However, Haugen and Baker (2010) point to their mean values for these variable coefficients are very large relative to their standard errors and argue that multicollinearity is clearly not a significant problem in their analysis.

[8]With unconstrained optimization, with 24 monthly observations and 1000 stocks, there is no unique solution. However, given the constraints provided above, unique solutions exist.

2007 period. Obviously, as Haugen and Baker conclude, transactions costs would have to be unrealistically extreme to significantly close the gap between the high and low expected return portfolios. Haugen and Baker (2010) conclude with the following findings: (1) measures of current profitability and cheapness are overwhelmingly significant in determining the structure of the cross-section of stock returns; (2) the statistical significance of risk is also overwhelming, but the payoff to risk has the wrong sign period after period. The riskiest stocks over measures including market beta, total return volatility, and residual volatility tend to have the lowest returns; (3) 1-year momentum pays off positively, and that last month's residual return and last quarter's total return pays off negatively; (4) strikingly, nearly all of the most significant factors over our total period are highly significant in our five sub-periods, and all have the same signs as they did in the total period; and (5) the ad hoc expected return factor model is very powerful in predicting the future relative returns on stocks.[9]

Stone and Guerard (2010) and Stone (2016) reestimated the Bloch et al. (1993) model for the 1967–2004 period and created portfolios using ranking on a forecasted return score and grouped securities into portfolios ordered on the basis of forecasted or predicted return score. This return cross section will almost certainly have a wide range of forecasted return values. However, each portfolio in the cross section will almost never have the same average values as a set of the control variables. To the extent values of return impact controls fluctuate randomly about their average value over the cross section, the variation is a source of noise (and possible loss of efficiency) in assessing the cross sectional dependency on the return forecast score on realized returns (and realized return related measures such as standard deviation). However, to the extent that a control is correlated with the return forecast score, the systematic variation in the control with the rank ordering will mean the impact of the control variable is mixed in with the dependency on the return forecast. Reassigning stocks to produce a new cross section with no portfolio-to-portfolio variation in the control variable means there is no differential impact on the return cross sections. In addition to possible improvements in efficiency, making all of the portfolios in a cross section have the same portfolio-average value for a correlated control variable means that we can isolate the dependency of realized returns on the return forecast from any dependency without any differential distortion from the correlated control

[9]Haugen and Baker (2010, p. 14) actually close their work with the following observations: "High-return stock deciles tend to be relatively large companies with low risk and they have positive market price momentum. The profitability of high-return stocks is good and getting better. The low-return counterparts to these stocks have the opposite profile. A rational investor would likely find the high-return profile very attractive and the low-return profile very scary. Given the evidence, and this evidence *will* be reproduced by others, the following conclusions are undeniable.

- The cross-sectional payoff to risk is highly negative.
- The longitudinal payoff to risk is highly positive.
- The most attractive stock portfolios have the highest expected returns.
- The scariest stock portfolios have the lowest expected returns.

The stock market is inefficient. "Case closed."

variable. To produce a cross-sectional match on any of the control variables, they reassigned stocks. For instance, if we were trying to make each portfolio in the cross section have the same average beta value, we could move a stock with an above-average beta value into a portfolio whose average beta value is below the population average. At the same time, we could shift a stock with a below-average beta value into the above average portfolio from the below-average portfolio.

Just to produce a match for each portfolio in the cross section on a single explanatory control variable such as beta clearly entails an immense number of possible reassignments of stocks across portfolios. Fortunately, we do not have to use trial-and-error switching of stocks between portfolios to find the best reassignment that produces a cross-sectional match on beta or any other control variable. This reassignment problem can be formulated as a mathematical assignment program (*MAP*). All fractile portfolios should have explanatory controls equal to their population average value. Given a cross section of fractile portfolios formed by rank-ordered grouping on the basis of predicted return, the objective of the assignment program was to transform this cross section of fractile portfolios into an associated control-matched cross section to optimize two complementary attributes of statistical power:

1. Preserving a wide range of well-ordered return forecasts.
2. Preserving within-portfolio homogeneity of forecasted return.

The four constraints are:

1. The portfolio average value of each control variable must equal the population mean.
2. The initial size (number of securities) of each portfolio must be preserved.
3. Each security must be fully assigned.
4. There can be no short sales.

The crucial constraints were the control matching restrictions. Preserving initial portfolio size and full use of each security are technical constraints that go with full use of the sample.

Stone and Guerard (2010) assessed month-to-month return cross sections in each of the 456 months of the 1968–2004 time period, and imposed progressively more complete sets of control variables in each month. Obtaining 15 or more control-matched cross sections in 456 month means solving more than 6700 optimization runs. Solving this many quadratic programs would be a computational challenge. However, just as one can approximate the mean-variance portfolio optimization of Markowitz (1952, 1956, 1959) by solving an associated linear programming (LP) approximation to the quadratic program [see for instance Stone (1973)], we can approximate the control-matching quadratic optimization by an associated LP objective function. The substance of the reassignment process is well understood by knowing input and output. The input was a cross section formed by ranking stocks into 30 fractile portfolios, which is the focus of most cross-sectional return analyses in past work on cross-sectional return dependencies. The output was a cross section of fractile portfolios that were matched on a specified set of controls variables. The MAP found an optimal reassignment of stocks that transformed the input rank-ordered cross section into a new cross section that is matched on the portfolio average values of each control variable.

Optimization arises in finding the particular reassignment that optimizes a trade-off between preserving the widest possible range of well-ordered portfolio values of forecasted return and also ensuring preservation of within-portfolio homogeneity of forecasted return. For the LP transformation used to obtain control-matched cross sections, the linear objective function is a trade-off between preserving as much as possible a measure of cross-portfolio range while minimizing the shifting of stocks away from the original rank-order portfolio. Even though we have used an LP to construct a control-matched cross section, any approximation is in the objective function and not in meeting any constraints. The constraints requiring that every portfolio in the cross section have an exact match on each control variable is met exactly in each month for every control variable in every portfolio.

Stone and Guerard (2010) used nonfinancial stocks during the 1962–2004 period to form portfolios in the CRSP-COMPUSTAT universe intersections that have been listed at least 5 years at the time of portfolio formation for which all return and all financial statement data are available. Because of the sparseness of the Compustat database in the 1964–1966 5-year start-up period required for variables such as 5 years sales growth, there are only 324 companies in the first forecast month, January 1968. From 1971 on, there are more than 900 companies in the forecast sample growing to more than 2000 companies by 1995. The fact that the sample size shows little growth from the 2003 companies in January 1995 to 2238 companies in January 2004 indicates that the large number of new IPOs from the mid-1990s on does not produce an increase in the number of sample companies. The fact that the Stone and Guerard (2010) sample does not exhibit the same growth as the cross time increase in publicly listed companies shows that the combination of data availability and minimum book value restrictions mean that we are primarily studying larger, more mature companies.

Conventional practice in cross sectional return dependency assessments has been to form deciles and more recently only quintiles. There are several reasons for using 30 (fractile) portfolios rather than forming deciles or even quintiles as done in some recent studies. Using a larger number pertains to the power-efficiency trade-off. First, grouping (averaging) tends to lose information, especially in the tails of the distribution while most of the efficiency benefits of measurement error and omitted variable diversification are accomplished with 20 or fewer stocks in a fractile. Second, to do regression fits to the return and standard deviation cross sections, more portfolio observations are clearly preferred to less with at least 20 being desirable. Third, Stone (2003) shows that very low levels of sample correlation coefficients are magnified nonlinearly when one groups by rank-ordering on any one variable. Fourth, given that one can group together adjacent portfolios in a control matched return cross section (as done here in looking at both the top quintile (top six portfolios grouped) and bottom quintile (bottom six portfolios grouped)), one should probably err on the side of too many portfolios in a cross section rather than too few.

Given a return forecast for each stock in the sample and a rank ordering into portfolio fractiles, Stone and Guerard (2010) had a return cross-section with no controls. They input this cross section to the MAP with all reasonable controls and obtain thereby a well-isolated cross sectional dependency realized return on return forecast score to the extent that there are no omitted controls. However, rather than going from no controls to a full set of controls in one step, they added add controls in a stepwise fashion, initially adding beta, size, and the book-price ratio, the Fama and French (1992) risk controls, and evaluated the combination of all three risk controls together. Additionally, they added their three tax controls, Leverage, Earnings Yield, and Dividend Yield, as a group and then as a group with the three Fama-French risk controls, and reported that these three tax controls have a significant impact after imposing the standard risk controls. Thus, from stepwise imposition of controls, Stone and Guerard (2010) reported that the three standard risk controls alone do not ensure a well-isolated return-forecast response subsurface. Since the book-price ratio is a model variable, its use as a control means that any contribution of BP to return performance is removed from the cross section. The key point here is that Stone and Guerard (2010) isolated return response from not only non-model control variables but also forecast model variables. Thus, it is possible to assess the relative contribution to both return and risk of individual forecast variables or forecast variable combinations by making them into controls.

Stone and Guerard (2010) used as a starting point for input to the MAP a cross-section of stocks rank-ordered on forecasted return score and partitioned into 30 fractile portfolios. While one could simply go from this input to a cross section matched on all the controls, a researcher can learn more about the impact of controls on the cross section by imposing the control restrictions in a stepwise fashion. The stepwise imposition of controls outlined here is a progressive exploration of how the cross section of realized returns depends jointly on the return forecast score used to form the cross section of fractile portfolios and other return impact variables that may be correlated or at least partially correlated with the return forecast score. The process of stepwise imposition of progressively more complete control sets is a process of moving toward a progressively more well-defined (well-isolated) return-forecast response subsurface. Stone and Guerard presented evidence to support the use of Beta, Size, Book-to-Price, Earnings Yield, Dividend Yield, and Leverage risk controls. The Stone and Guerard (2010) conclusions were:

1. The long run average realized returns had a significant dependency upon forecast return score, especially for control sets F and beyond.
2. The linear coefficient on standard deviation were generally negative and insignificant when there was just a linear standard deviation term in the cross sectional regression. When there was also a quadratic term in the regression, the coefficient on the linear term becomes negative and significant. Either an insignificant or negative coefficient on standard deviation suggested that the return realizations were not explained by increasing risk as measured by realized standard deviation.
3. The fact that the skewness coefficient had very little incremental explanatory value beyond the forecast return score suggested that the realization of greater

right skewness with increasing return score did not explain the apparently significant potential for realizing alpha performance from utilization of the forecast.

Given that investors generally like and presumably correctly price right skewness, it is puzzling that realization of significant positive skewness was seemingly not priced in the realized return response to the forecast score.

Because standard deviation does not explain most of the performance potential inherent in the dependency of long run realized returns on return forecast score and given that there is additional positive skewness, the evidence for a positive alpha performance on the return response subsurface was even stronger with the positive skewness than would be the case if there were just this magnitude of returns without any positive skewness in the high return rank portfolios. Overall, the regressions summarized indicated that for the cross sections for which the return forecast as well isolated from other return impact variables, the significant long run realized return dependency on the forecast was primarily an alpha performance potential rather than a systematic risk effect.

Guerard, Gultekin, and Xu (2012) extended a stock selection model originally developed and estimated in Bloch et al. (1993), model, adding a Brush (2001, 2007)-based price momentum variable, taking the price at time $t-1$ divided by the price 12 months ago, $t-12$, denoted PM, and the consensus (I/B/E/S) analysts' earnings forecasts and analysts' revisions composite variable, CTEF, to the stock selection model.[10] Other momentum strategies can be found in Korajczyk and Sadka (2004), and Fama and French (2008). Guerard et al. (2012) referred to the stock selection model as a United States Expected Returns (USER) Model. We can estimate an expanded stock selection model to use as an input of expected returns in an optimization analysis. The universe for all analysis consists of all securities on Wharton Research Data Services (WRDS) platform from which we download the I/B/E/S database, and the Global Compustat databases. The I/B/E/S database contains consensus analysts' earnings per share forecast data and the Global Compustat database contains fundamental data, i.e.; the earnings, book value, cash flow, depreciation, and sales data, used in this analysis for the January 1990 to December 2010 time period. The information coefficient, IC, is estimated as the slope of a regression line in which ranked subsequent returns are expressed as a function of the ranked strategy, at a particular point of time. The high fundamental variables, earnings, book value, cash flow, and sales produce higher ICs in the Global universe

[10]Guerard (2012) decomposed the MQ variable into (1) price momentum, (2) the consensus analysts' forecasts efficiency variable, CIBF, which itself is composed of forecasted earnings yield, *EP*, revisions, EREV, and direction of revisions, *EB*, identified as breadth, Wheeler (1994), and (3) the stock standard deviation, identified as a variable with predictive power regarding the stock price-earnings multiple. Guerard reported that the consensus analysts' forecast variable dominated analysts' forecasted earnings yield, as measured by I/B/E/S 1-year-ahead forecasted earnings yield, FEP, revisions, and breadth. Guerard reported domestic (US) evidence that the predicted earnings yield is incorporated into the stock price through the earnings yield risk index. Moreover, CIBF dominates the historic low price-to-earnings effect, or high earnings-to-price, *PE*.

than in the US universe where USER was estimated. Moreover, analysts' 1-year-ahead and 2-year ahead revisions, RV1 and RV2, respectively, were much lower in Global markets, than US market. Breadth, BR, and forecasted earnings yields, FEP, were positive but less than in the US market. The consensus earnings forecasting variable, CTEF, and the price momentum variable, PM, dominate the composite model. We use the CTEF model without addressing the issue of whether analysts' or rational.[11] Analysts' forecast revisions, breadth, and yields are more effective than analysts' forecasts in portfolio construction and management. The PM variable

[11]Cragg and Malkiel (1968) created a database of five forecasters of long-term earnings forecasts for 185 companies in 1962 and 1963. These five forecast firms included two New York City banks (trust departments), an investment banker, a mutual fund manager, and the final firm was a broker and an investment advisor. The Cragg and Malkiel (1968) forecasts were 5-year average annual growth rates. The earnings forecasts were highly correlated with one another, the highest paired correlation was 0.889 (in 1962) and the lowest paired correlation was 0.450 (in 1963) with most correlations exceeding 0.7. They calculated used the Thiel Inequality Coefficient (1966) to measure the efficiency of the financial forecasts and found that the correlations of predicted and realized earnings growth were low, although most were statistically greater than zero. The TICs were large, according to Cragg and Malkiel (1968), although they were less than 1.0 (showing better than no-change forecasting). The TICS were lower (better) within sectors; the forecasts in electronics and electric utility firms were best and foods and oils were the worst firms to forecast earnings growth. Elton and Gruber (1972) built upon the Cragg and Malkiel study and found similar results. That is, a simple exponentially weighted moving average was a better forecasting model of annual earnings than additive or multiplicative exponential smoothing models with trend or regression models using time as an independent variable. Indeed, a very good model was a naïve model, which assumed a no-change in annual earnings per shares with the exceptional of the prior change had occurred in earnings. One can clearly see the random walk with drift concept of earnings in the Elton and Gruber (1972). Elton and Gruber compared the naïve and time series forecasts to three financial service firms, and found for their 180 firm sample that two of the three firms were better forecasters than the naïve models. Elton, Gruber, and Gultekin (1981) build upon the Cragg and Malkiel (1968) and Elton and Gruber (1972) results and create an earnings forecasting database that evolves to include over 16,000 companies, the Institutional Brokerage Estimation Services, Inc. (I/B/E/S). Elton et al. (1981) find than earnings revisions, more than the earnings forecasts, determine the securities that will outperform the market. Found the I/B/E/S consensus forecasts were not statistically different than random walk with drift time series forecasts for 648 firms during the 1982–1985 period. Lim (2001), using the I/B/E/S Detailed database from 1984 to December 1996, found forecast bias was associated with small and more volatile stocks, experienced poor past stock returns, and had prior negative earnings surprises. Moreover, Lim (2001) found that relative bias was negative associated with the size of the number of analysts in the brokerage firm. That is, smaller firms with fewer analysts, often with more stale data, produced more optimistic forecasts. Keane and Runkle (1998) found during the 1983–1991 period that analysts' forecasts were rational, once discretionary special charges are removed. The Keane and Runkle (1998) study is one of the very few studies finding rationality of analysts' forecasts; most find analysts are optimistic. Further work by Wheeler (1994) will find that firms where analysts agree with the direction of earnings revisions, denoted breadth, will outperform stocks with lesser agreement of earnings revisions. Guerard et al. (1997) combined the work of Elton et al. (1981) and Wheeler (1994) to create a better earnings forecasting model, CTEF, which we use in this analysis. The CTEF variable continues to produce statistically significant excess return in backtest and in identifying real-time security mispricing, see Guerard (2012). See Brown (1999, 2008) and Ramnath, Rock, and Shane (2008) for an extensive review of financial analysts' forecasting efficiences.

is calculated as the price 1-month ago divided by the price 12 months ago, in the tradition of Brush and Boles (1983).[12]

The stock selection model estimated in this chapter, denoted as United States Expected Returns, USER, and Global Expected Returns, GLER, is:

$$\begin{aligned} \mathrm{TR}_{t+1} &= a_0 + a_1 EP_t + a_2 BP_t + a_3 CP_t + a_4 SP_t + a_5 REP_t \\ &\quad + a_6 RBP_\mathrm{t} + a_7 RCP_t + a_8 RSP_t + a_9 CTEF_t + a_{10} PM_t + e_t \end{aligned} \tag{4.3}$$

where:

$$EP = \frac{[\text{earnings per share}]}{[\text{price per share}]} = \text{earnings} - \text{price ratio};$$

$$BP = \frac{[\text{book value per share}]}{[\text{price per share}]} = \text{book} - \text{price ratio};$$

$$CP = \frac{[\text{cash flow per share}]}{[\text{price per share}]} = \text{cash flow} - \text{price ratio};$$

$$SP = \frac{[\text{net sales per share}]}{[\text{price per share}]} = \text{sales} - \text{price ratio};$$

$$REP = \frac{[\text{current } EP \text{ ratio}]}{[\text{average } EP \text{ ratio over the past 5 years}]};$$

$$RBP = \frac{[\text{current } BP \text{ ratio}]}{[\text{average } BP \text{ ratio over the past 5 years}]};$$

$$RCP = \frac{[\text{current } CP \text{ ratio}]}{[\text{average } CP \text{ } r\text{atio over the past 5 years}]};$$

$$RSP = \frac{[\text{current } SP \text{ ratio}]}{[\text{average } SP \text{ ratio over the past 5 years}]};$$

$\mathrm{CTEF} = \text{consensus earnings-per-share } eps \text{ forecast, revisions, and breadth};$
$\mathrm{PM} = \text{Price Momentum};$

and

$$e = \text{randomly distributed error term.}$$

The GLER model is estimated using a weighted latent root regression, WLRR, analysis on Eq. (4.3) to identify variables statistically significant at the 10 % level; uses the normalized coefficients as weights; and averages the variable weights over

[12] The ICs on the analysts' forecast variable, CTEF, and price momentum variable, PM, were lower than in the US universe, reported in Guerard et al. (2012).

the past 12 months. The weighting technique used was the Beaton and Tukey (1974) bisquare procedure for robust regression. The reader is referred to Gunst, Webster, and Mason (1976) and Gunst and Mason for LRR analysis and Rousseeuw and LeRoy (1987), Yohai (1987), Yohai, Stahel, and Zamar (1991), and Maronna, Martin, and Yojai (2006) for more complete discussions of robust regression. The 12-month smoothing is consistent with the four-quarter smoothing in Bloch et al. (1993). While *EP* and *BP* variables are significant in explaining returns, the majority of the forecast performance is attributable to other model variables, namely the relative earnings-to-price, relative cash-to-price, relative sales-to-price, price momentum, and earnings forecast variables. The CTEF and PM variables account for 48 % of the model average weights, slightly higher than the two variables combining for 44 % of the weights in the USER Model.

The Time-Average Value of GLER Estimated Coefficients:

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0.048	0.069	0.044	0.047	0.050	0.032	0.039	0.086	0.216	0.257

In terms of information coefficients, ICs, the use of the WLRR procedure produces a virtually identical IC for the models during the 1990–2010 time period, 0.042, versus the equally weighted IC of 0.043. The GLER model, has compared to the USER Model in Guerard et al. (2012) has approximately the same ICs. The IC test of statistical significance can be referred to as a Level I test (Table 4.1).

There is strong support for fundamental variables (particularly earnings and cash flow), earnings expectations variables, and the momentum variable. An objective examination of the reported ICs leads one to identify CTEF, PM, *EP*, and *CP* as leading variables for inclusion in stock selection models.

4.3 Constructing Mean-Variance Efficient Portfolios

The origin of modern finance in this context must be traced to the work of Markowitz (1952, 1956, 1959). The conceptual framework is based on the work of von Neumann and Morgenstern (1944) who pioneered the view that choice under uncertainty may be based on expected utility. As initially formulated by Markowitz this involved the *maximization of portfolio returns given a variance constraint*. In practice the standard deviation was substituted for the variance to eliminate dependence on the units of measurement, e.g., dollars versus thousands of dollars. If one solves this problem one finds that at the optimum the Lagrange multiplier, λ, equals what came to be known as the Sharpe ratio. The interpretation of λ in other areas of economics is that of a “shadow price” and measures the extent to which the function to be maximized, here expected portfolio returns, will change by relaxing the constraint, here risk, at the optimum. Thus it may have the interpretation as the marginal return on risk, or the reward (or price) of risk at the optimum. This has

Table 4.1 Global composite model component ICs 1/1990–9/2009

Variable	IC
EP	0.048
BP	0.019
CP	0.042
SP	0.008
DP	0.058
RV1	0.011
RV2	0.019
BR1	0.026
BR2	0.024
FEP1	0.034
FEP2	0.029
CTEF	0.023
PM	0.022
EWC	0.043
GLER	0.042

led to the widespread belief that the portfolio manager should seek to maximize the portfolio geometric mean (GM) and Sharpe ratio (ShR), as put forth in Latane (1959) and Markowitz (1959 and 1976).[13] However, as formulated in Markowitz the portfolio manager seeks to identify the efficient frontier, the point at which the portfolio return is maximized for a given level of risk, or, equivalently, portfolio risk is minimized for a given level of portfolio return. The portfolio expected return, denoted by $E(R_p)$, is calculated by taking the sum of the security weight multiplied by their respective expected return. The portfolio standard deviation is the sum of the weighted security covariance.

$$E\left(R_p\right) = \sum_{i=1}^{N} w_i E\left(R_i\right) \tag{4.4}$$

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij} \tag{4.5}$$

where N is the number of candidate securities, w_i is the weight for security i such that $\sum_{i=1}^{N} w_i = 1$ indicating that the portfolio is fully invested, and $E(R_i)$ is the expected return for security i.

The Markowitz framework measures risk as the portfolio standard deviation, a measure of dispersion or total risk. One seeks to minimize risk, as measured by the covariance matrix in the Markowitz framework, holding constant expected returns. Elton, Gruber, Brown, and Goetzman (2007) in fact proposed what they conceived

[13] See Markowitz (1959), Chapter 9.

to be an equivalent formulation of the traditional Markowitz mean-variance problem as a maximization problem, i.e., to maximize

$$\theta = \frac{E\left(R_p\right) - R_{\mathrm{F}}}{\sigma_p} \tag{4.6}$$

where R_{F} is the risk-free rate (typically measured by the 90-day Treasury bill rate). A little reflection, however, will show that it is not; in fact these are two conceptually separate problems.

Implicit in the development of the Capital Asset Pricing Model, CAPM, by Sharpe (1964), Lintner (1965a, 1965b), and Mossin (1966) is that investors are compensated for bearing systematic or market risk, not total risk. Systematic risk is measured by a stock's beta. Beta is the slope of the market model in which the stock return is regressed as a function of the market return.[14] Sharpe (1963) proposed what he refers to as the diagonal model to simplify the computations in constructing portfolios. An investor is not compensated for bearing risk that may be diversified away from the portfolio. The reader is referred to Rudd and Clasing (1982) and Stone (1970) for early treatments of the risk-return trade-off and Markowitz (2013) for the sixty year perspective.

The CAPM holds that the return to a security is a function of the security's beta.

$$R_{jt} = R_{\mathrm{F}} + \beta_j \left[E\left(R_{\mathrm{M}t}\right) - R_{\mathrm{F}}\right] + e_j \tag{4.7}$$

where

$$R_{jt} = \text{expected security return at time } t;$$

$$E\left(R_{\mathrm{M}t}\right) = \text{expected return on the market at time } t;$$

$$R_{\mathrm{F}} = \text{risk-free rate;}$$

$$\beta_j = \text{security beta; and}$$

$$e_j = \text{randomly distributed error term.}$$

An examination of the CAPM beta, its measure of systematic risk, from the Capital Market Line equilibrium condition follows.

[14] Harry Markowitz reminds readers that he discussed the possibility of looking at security returns relative to index returns in Chapter 4, footnote 1, page 100, of *Portfolio Selection* (1959).

$$\beta_j = \frac{\text{Cov}\left(R_j, R_\text{M}\right)}{\text{Var}\left(R_\text{M}\right)} \tag{4.8}$$

The difficulty of measuring beta and its corresponding Security Market Line, SML, gave rise to extra-market measures of risk found in the work of Farrell (1974, 1997), Rosenberg (1974), Ross (1976), Ross and Roll (1980), Dhrymes, Friend, and Gultekin (1984) and Dhrymes, Friend, Gultekin, and Gultekin (1985).[15] The fundamentally based domestic Barra risk model was developed in a series of studies by Rosenberg and Marathe (1979), Rudd and Rosenberg (1980) and thoroughly discussed in Rudd and Clasing (1982) and Grinhold and Kahn (1999), Connor and Korajczyk (1988, 1993, 1995, and 2010), and Connor, Goldberg, and Korajczyk (2010).

Guerard (2012) demonstrated the effectiveness of the Ross Arbitrage Pricing Model (APT) and Sungard APT systems in portfolio construction and management. Let us review the APT approach to portfolio construction. The estimation of security weights, w, in a portfolio is the primary calculation of Markowitz's portfolio management approach. The issue of security weights will be now considered from a different perspective. The security weight is the proportion of the portfolio's market value invested in the individual security.

$$w_s = \frac{\text{MV}_s}{\text{MV}_p} \tag{4.9}$$

where w_s = portfolio weight in security s, MV_s = value of security s within the portfolio and MV_p = the total market value of portfolio.

The active weight of the security is calculated by subtracting the security weight in the (index) benchmark, b, from the security weight in the portfolio, p

$$w_{s,p} - w_{s,b} \tag{4.10}$$

Markowitz analysis (1952, 1959) and its efficient frontier minimized risk for a given level of return. Blin and Bender created APT, Advanced Portfolio Technologies, Analytics Guide (2011), which built upon the mathematical foundations of their APT system, published in Blin, Bender, and Guerard (1997). The following analysis draws upon the APT analytics. Volatility can be broken down into systematic and specific risk:

$$\sigma_p^2 = \sigma_{\beta p}^2 + \sigma_{\varepsilon p}^2 \tag{4.11}$$

where σ_p = Total Portfolio Volatility, $\sigma_{\beta p}$ = Systematic Portfolio Volatility and $\sigma_{\varepsilon p}$ = Specific Portfolio Volatility. Blin and Bender created a multifactor risk model within their APT risk model based on forecast volatility.

[15] See Chapter 2 of Guerard (2010) for a history of multi-index and multifactor risk models.

$$\sigma_p = \sqrt{52\left(\sum_{c=1}^{c}\sum_{i=1}^{s} w_i\beta_{i,c}\right)^2 + \sum_{i=1}^{s}\mathbf{w}_i^2\varepsilon_{i,\text{week}}^2} \tag{4.12}$$

where $\sigma_p =$ forecast volatility of annual portfolio return,
$C =$ number of statistical components in the risk model,
$w_i =$ Portfolio weight in security i,
$\beta_{i,c} =$ loading (beta) of security i on risk component c,
$\varepsilon_{i,\text{week}} =$ weekly specific volatility of security i.

Portfolio-specific volatility is a forecast of the annualized standard deviation associated with each security's specific return.

$$\sigma_{\varepsilon p} = \sqrt{52\sum_{i=1}^{s}\mathbf{w}_i^2\varepsilon_{i,\text{week}}^2} \tag{4.13}$$

Tracking error is a measure of volatility applied to the active return of funds (or portfolio strategies) indexed against a benchmark, which is often an index. Portfolio tracking error is defined as the standard deviation of the portfolio return less the benchmark return over 1 year.

$$\sigma_{\text{te}} = \sqrt{E\left(\left((r_p - r_b) - E\left(r_p - r_b\right)\right)^2\right)} \tag{4.14}$$

where $\sigma_{\text{te}} =$ annualized tracking error,
$r_p =$ actual portfolio annual return,
$r_b =$ actual benchmark annual return.

Systematic tracking error of a portfolio is a forecast of the portfolio's active annual return as a function of the securities' returns associated with APT risk model components. Portfolio-specific tracking error can be written as a forecast of the annual portfolio active return associated with each security's specific behavior.

$$\sigma_{\varepsilon\text{te}} = \sqrt{52\sum_{i=1}^{S}\left(\mathbf{w}_{i,p} - \mathbf{w}_{i,b}\right)^2\varepsilon_{i,\text{week}}^2} \tag{4.15}$$

The APT calculated portfolio error versus a benchmark is:

$$\sigma_{p-b} = \sqrt{52\left(w_p - w_b\right)^{\mathrm{T}}\left(b^{\mathrm{T}}b + \varepsilon^{\mathrm{T}}\varepsilon\right)\left(w_p - w_b\right)} \tag{4.16}$$

where $\sigma_{p-b} =$ forecast tracking error
$b =$ A $(N_c \times N_s)$ matrix of component loadings; N_c components in the model and N_s securities in the portfolio,
$\varepsilon =$ A diagonal matrix $(N\,(N_s \times N_s))$ of the specific loadings,

$w_p - w_s =$ The ($N_s -$ dimensional) vector security weights.

$$\sigma_{p-b} = \sqrt{52\left(\sum_{c=1}^{N_c}\left(\sum_{s=1}^{N_s}\left(w_{sp} - w_{sb}\right)b_{sc}\right)^2 + \sum_{s=1}^{N_s}\left(w_{sp} - w_{sb}\right)^2\varepsilon_s^2\right)} \tag{4.17}$$

where $w_{sp} - w_{sb} =$ portfolio active weights in security s,

$b_{sc} =$ the loading of security s on component c,

$\varepsilon_s =$ weekly specific volatility of security s.

The marginal tracking error is a measure of the sensitivity of the tracking error of an active portfolio to changes in active weight of a specific security.

$$\partial_s\left[\sigma_{p-b}\right] = \frac{\partial\sigma_{p-b}}{\partial w_{s,p-b}} \tag{4.18}$$

where $w_{s,p-b} =$ active portfolio weights in security s.

The APT calculated contribution-to-risk of a security is:

$$\Delta\left[\sigma_p\right] = \frac{\left(b^{\mathrm{T}}b + \varepsilon^{\mathrm{T}}\varepsilon\right)w}{\sqrt{52w^{\mathrm{T}}\left(b^{\mathrm{T}}b + \varepsilon^{\mathrm{T}}\varepsilon\right)w}} \tag{4.19}$$

where $\Delta[\sigma_p] =$ A ($N_s -$ dimensional) vector of contribution to volatility for securities in the portfolio with N_s securities.

The portfolio Value-at-Risk (VaR) is a measure of the distribution of expected outcomes. If one is concerned with a 95 % confidence level, α, and a 30-day time horizon, then the 95 %, 30-day VaR of the portfolio is the minimum we would expect to lose 95 % over a 30-day period

$$P\left[V_T < V_0 - v\left(\alpha, T\right)\right] = 1 - \alpha \tag{4.20}$$

where $v(\alpha, T) =$ VaR at the confidence interval α for time T,

$V_T =$ portfolio value at time T,

$\alpha =$ required confidence level.

If the portfolio returns are normally distributed,

$$v_s\left(\alpha, T\right) = V_0 - E\left[V_T\right] - V_0\phi^{-1}\left(\alpha\right)\sigma_p^T \tag{4.21}$$

where $v_s(\alpha, T) =$ Gaussian VaR,

$\sigma_p^T =$ forecast portfolio volatility,

$\phi^{-1}(\alpha) =$ inverse cumulative normal distributed function.

The total VaR is:

$$v_c(\alpha, T) = V_0 \sqrt{\frac{\alpha}{1-\alpha}\left(\sigma_{\beta_p}^T + \left(\phi^{-1}(\alpha)\,\sigma_{\varepsilon_p}^T\right)^2\right)} \tag{4.22}$$

The Tracking Error at Risk (TaR) is a measure of portfolio risk estimating the magnitude that the portfolio may deviate from the benchmark over time *dt* is the maximum deviation of a portfolio return over the time horizon T at a given confidence level α.

$$r_{p-b}(\alpha, T) = \sqrt{\left(\frac{1}{\sqrt{1-\alpha}}\sigma_{\beta_{p-b}}^T + \phi^{-1}\left(\frac{1+\alpha}{2}\right)\sigma_{\varepsilon_{p-b}}^2\right.} \tag{4.23}$$

Blin, Bender, and Guerard (1997) used an estimated 20-factor beta model of covariances based on 3.5 years of weekly stock returns data. The Blin and Bender Arbitrage Pricing Theory (APT) model followed the Ross factor modeling theory, but Blin and Bender estimated betas from at least 20 orthogonal factors. Empirical support is reported in Guerard (2012), for the application of mean-variance, enhanced index tracking and tracking error at risk optimization techniques. It is well known that as one raises the portfolio lambda, the expected return of a portfolio rises and the number of securities in the optimal portfolios fall, see Blin, Bender, and Guerard (1997). Lambda, a measure of risk-aversion, the inverse of the risk-aversion acceptance level of the Barra system, is a decision variable to determine the optimal number of securities in a portfolio. The Blin and Bender TaR optimization procedure allows a manager to use fewer stocks in his or her portfolios than a traditional mean-variance optimization technique manager for a given lambda. In spite of the Markowitz Mean-Variance portfolio construction and management analysis being six decades old, it does very well in maximizing the Sharpe Ratio, Geometric Mean, and Information Ratio relative to newer approaches.

4.4 Evaluation of Portfolio Performance: Origins

Earlier we remarked on the interpretation of the Lagrange multiplier, λ, and we noted that at the optimum it denotes the ratio of expected excess returns to risk usually denoted by the standard deviation; the latter is of course an *ex ante measure* of the worth or reward (in terms of expected returns) per "unit" of risk (standard deviation or variance) undertaken, at the optimum.

One may think of this construct *ex post* as a measure or index of the performance of a portfolio manager. Although this problem was not approached initially with this perspective, an important paper by Treynor (1965) gave impetus for the research of such issues. He proposed the index

$$\text{TI} = \text{excess returns/a slope}$$

Because the concept is represented by a graph we will first give an algebraic interpretation relabeling what Treynor calls the return to fixed assets, as the risk free rate r_0, for simplicity of reference. In his framework there is a characteristic line for the typical fund, which I shall interpret as the expected value of the relationship

$$r_{p(i)} = a + br_{\mathrm{m}} + u_p(i),$$

so that the expected return for a typical portfolio (perhaps only in a certain class) is $a + br_{\mathrm{m}}$, r_{m} being the market rate. If a given portfolio (ex-post) is characterized by a pair $(r_p(i), r_{\mathrm{m}^*})$ according to Treynor's index, it would be awarded the rank ρ which is defined by

$$\rho = r_{\mathrm{m}*} - (r_0 - a)/b;$$

this, evidently, was not deemed sufficiently informative so it was rewritten (by Treynor) as[16]

$$\rho = r_{\mathrm{m}*} - \left(r_p(i)\text{–}r_0\right)/\eta, \quad \eta = \left(r_p(i) - r_0\right)/\left(r_{\mathrm{m}*} - \rho\right),$$

Notice that η is the *tangent of the angle made by the intersection of the horizontal line* beginning on the vertical axis at point r_0 and the line $a + b\ r_{\mathrm{m}}$, (denoted in the paper by σ, although it is not a standard deviation). The intuitive sense of Treynor's measure is this: given the characteristic line, no great skill is required to operate only with fixed (risk free) assets earning r_0; in the universe of risky funds with the characteristic line $a + b\ r_{\mathrm{m}}$, this (return) would correspond to the market rate $(r_0 - a)/b$; if a fund manager attains $(r_p(i), r_{\mathrm{m}^*})$, then what is due to his skill is $r_p(i) - r_0$, i.e., the excess return; moreover, the steeper the characteristic line the easier it is to attain this prespecified difference; hence it is normalized by the slope. To further clarify the concept, note that from the characteristic line this excess return $(r_p(i) - r_0)$ would correspond to a certain market rate, say $r_{\mathrm{m}^{**}}$. Hence a manager's rank is given by $r_{\mathrm{m}^*} - r_{\mathrm{m}^{**}}$, so that the manager who had the "smaller help from the market", in the form of a lower $r_{\mathrm{m}^{**}}$, to gain a given amount of excess return, $r_p(i) - r_0$, gets a higher rating.[17] Thus, the Treynor index, in principle, has nothing to do with the standard deviation of the portfolio, which in fact cannot be determined given the information in the paper.

In a subsequent paper, joint with K. K. Mazuy (1966), a similar but somewhat different question is asked, viz. whether fund managers can accurately "guess" the direction of the market. This paper examines the performance of 57 mutual funds over the period 1953–1962. The main tool is again the "market or characteristic

[16]Note that when symbols are assigned their proper meaning the first equation below simply states $\rho = \rho$.

[17]Although not perhaps in the Treynor spirit a more useful index might be $\delta = (r_p(i) - r_0)/(a + b\ r_{\mathrm{m}}^*)$, the ratio of excess to market returns, in lieu of the original form $\mathrm{TI} = (r_p(i) - r_0)/b$.

line"; here two regimes are posited: either regime has its own characteristic line. It is argued that if a manager successfully navigates between the two regimes, i.e., he more or less correctly "reads" the market then the fund's characteristic line (or at least its scattergram) should exhibit curvature. *He finds this to be approximately true only for one fund*. The reporting requirements of the time do not permit the precise interpretation of the F statistic on which they base this finding; I take it to mean that for only one fund did they find a "statistically significant" quadratic term in the characteristic line. But then with seven degrees of freedom this may be asking for more than the accuracy of the data may support.

Their conclusion, however, contains very wise advice to portfolio managers and those who rely on their work.

Another approach to the same problem is given in Sharpe (1966); it involves using the inverse of the coefficient of variation (mean over the standard deviation) as a means of evaluating fund performance. The coefficient of variation is well known from probability theory, where it is used as an index of dispersion of the distribution of a random variable; its inverse, when based on estimators of the mean and standard deviation and *properly normalized*, is t-distributed if the underlying distribution is normal or is *asymptotically normally distributed* if the observations underlying the estimators obey certain conditions which justify the application of one of the classical central limit theorems. The inverse of the coefficient of variation in the finance literature is known as the *Sharpe ratio*[18] and was introduced by Sharpe (1966) in connection with the evaluation of mutual fund performance; it is also referred to as the reward variability ratio.

Despite its relation to the Lagrange multiplier and its affinity with the t-statistic, which is used to test statistical hypotheses, it is not clear why the Sharpe ratio is an appropriate index for judging and ranking the performance of portfolios. For one thing, a portfolio over a decade has no fixed composition, i.e., the proportions devoted to its components vary over time. Thus, even if *all individual betas remain reasonably fixed over the period*, *the parameters of the expected (mean) returns and risk (standard deviation) vary over time because the components of the portfolio vary*. Thus, a procedure that mimics what we do when we estimate parameters we assert to be fixed over the sample period, has nothing particularly to recommend it in this context. For another a portfolio consisting mostly of near risk free assets, such as money market like assets may have a high Sharpe ratio because even though its numerator is very small the denominator (risk, standard deviation) may be nearly zero!

Probably, to measure performance we need to display *both excess returns and risk (standard deviation)*, so that within each risk class they can be unambiguously ranked.

[18]In Sharpe (1966) it is somewhat incorrectly stated that the Sharpe ratio is equivalent to the Treynor ratio. As is seen from the discussion above noting in particular the modification in footnote 17, this is not true; the Treynor ratio as modified in that footnote is merely the ratio of excess returns divided by market returns, not the standard deviation.

Of course, given that all such quantities are estimated from empirical data they are random entities estimating the underlying parameters and are thus subject to a margin of error.

While the computational aspects of both indices are very well defined, the strength or robustness of conclusions derived there from is not evident; for example if a fund exhibits the Treynor value τ and another $\tau + c$ and c is relatively small and positive, is fund 2 better than fund 1? We do not have any sensitivity basis and we certainly do not have a distribution for the indices. Thus all such rankings are essentially judgmental and in the case of the Sharpe ratio can only be supported by vague appeals to the *t*-statistic like feature of the index. It is clear that the magnitudes of these indices are related to the astuteness of portfolio managers in choosing the elements of the vector $\mathbf{w}$ in response to market conditions. But this *ipso facto* could preclude the consideration of the Sharpe ratio as an estimator of an underlying parameter of the portfolio, since the latter's parameters are subject to unspecified changes over the period being evaluated.

4.5 Portfolio Simulation Results with the USER and GLER Models

Let us briefly review the Guerard et al. (2012) USER Model simulation to provide a baseline for comparison with the WRDS GLER Model in global markets during the 1999–2010 period. The portfolio returns of the USER model with MVTar and a lambda of 200 are shown in Table 4.2. We report Axioma attribution that the USER Model produced 10.7 % annual active return, a result consistent with the Barra attribution results reported in Guerard et al. (2012).[19] The active return is derived from a specific return of 16.3 % annually. The active return has an Information Ratio of 1.12 and a *t*-statistic of 3.68. Thus, the USER Model is effective because of "bottom-up" stock selection. The APT-derived portfolios have active exposures to Momentum, Value, and Size (0.49, 0.43, and −1.05, respectively) that are "priced" by the market to produce statistical significant portfolio factor returns with *t*-statistics of 4.07, 4.24, and 2.9, respectively.[20] The USER Model portfolios have positive exposure to value and momentum and smaller stocks are purchased.

[19]The authors are indebted to Vishhu Anand, of Axioma, who ran the Axioma attribution analysis based on the Axioma Fundamental Risk Model.

[20]Readers may question the use of a 12-year backtesting period. The USER model was tested with the Barra USE3 Model for the 1980–2009 period and asset selection of 449 basis points, annualized, is reported. The *t*-value on the USER variable is 4.60, which is highly statistically significant. Stone and Guerard (2010b) found good stock selection returns in the USA and Japan, 1980–2005, using similar models.

Table 4.2 Characteristics of APT-estimated risk constrained portfolios (WRDS GLER Model, 1999–2009)

	Portfolio performance criteria			
	NoRC	MRC	SRC	ACWG benchmark
Geometric mean	12.68 %	8.63 %	7.65 %	1.79 %
Information ratio	1.00	0.74	0.68	
Sharpe ratio	0.47	0.28	0.24	−0.036

Where NoRC = No Risk Controls GLER L = 200, MRC = Moderate Risk Controls, SRC = Strong Risk Controls, STD = Portfolio Standard Deviation

We experimented with the Wormald and van der Merve (2012) risk control conditions. We find that the No Risk (NoRC) Risk Control condition produced higher Information Ratios and Geometric Means than the Strong Risk Control (SRC) or Mild Risk Control (MRC) in our work, particularly in the global market. The WRDS GLER Model in global markets during the 1999–2009 period, created with Global Compustat and I/B/E/S data produces highly significant Geometric Means and Information Ratios, see Table 4.3. We experimented with the Wormald and van der Merve (2012) risk control conditions.

In the world of business, one does not access academic databases annually, or even quarterly. Most industry analysis uses FactSet database and the Thomson Financial (I/B/E/S) earnings forecasting database. We can estimate Eq. (4.3) for all securities on the Thomson Financial and FactSet databases, some 46,550 firms in December 2011. We estimated the GLER Model upon the FactSet universe, denoted FSGLER, of 46,000 stocks for the 1990–2011 period. We restricted the simulations to securities covered by at least two I/B/E/S analysts, a world of business simulation condition of McKinley Capital Management, MCM, using the Wormald and van der Merve (2012) risk control conditions. We find that the No Risk (NoRC) Risk Control condition produced higher Information Ratios and Geometric Means than the Strong Risk Control (SRC) or Mild Risk Control (MRC). These results are shown in Table 4.4.

We have established that the NoRC Model was effective in the FSGLER universe. The portfolio returns of the FSGLER NoRC model with a lambda of 200 are shown in Table 4.5. We report that the GLER Model produced 13.0 % annual active return, a result consistent with the USER attribution results. The active return is derived from a specific return of 16.3 % annually. The active return has an Information Ratio of 1.23 and a t-statistic of 4.44.

The GLER Model also is effective because of stock selection. The APT-derived portfolios have active exposures to Momentum, Value, and Size (0.36, 0.51, and −.96, respectively) that are "priced" by the market to produce statistical significant

Table 4.3 WRDS USER Model Barra USE3 attribution analysis (February 2000–December 2013)

Attribution report				
Annualized contributions to total return				
Source of return	Contribution (% return)	Risk (% Std Dev)	Info ratio	T-Stat
1 Risk free	1.97			
2 Total benchmark	2.05	18.30		
3 Expected active	0.34			
4 Market timing	0.33	2.62	0.19	0.70
5 Risk indices	5.57	5.41	1.00	3.71
6 Industries	0.51	3.84	0.18	0.68
7 Asset selection	5.02	4.50	1.07	3.99
8 Trading				
9 Transaction cost	−2.04			
10 Total exceptional active[$4+\cdots+9$]	9.39	8.67	1.11	4.12
11 Total active [$3+10$]	9.73	8.67	1.14	4.26
12 Total managed [$2+11$]	11.78	21.84		

(continued)

Table 4.3 (continued)

Annualized contributions to risk index return

		Contribution (% return)					
		Average	Variation	Total	Total		
Source of return	Average active exposure	[1]	[2]	[1 + 2]	Risk (% Std Dev)	Info ratio	*T*-Stat
Volatility	0.40	0.05	0.00	0.05	2.07	0.04	0.14
Momentum	0.01	−0.02	0.17	0.15	1.08	0.15	0.55
Size	−1.52	5.03	0.46	5.49	4.81	1.08	4.05
Size nonlinearity	−0.75	−1.07	0.22	−0.85	2.65	−0.30	−1.11
Trading activity	−0.06	0.02	0.14	0.16	0.25	0.60	2.25
Growth	−0.24	0.11	0.03	0.14	0.54	0.26	0.97
Earnings yield	0.16	0.57	0.65	1.22	0.86	1.26	4.70
Value	0.57	−0.15	0.01	−0.14	1.15	−0.08	−0.30
Earnings variation	0.36	−0.27	−0.04	−0.31	0.69	−0.36	−1.34
Leverage	0.39	−0.25	0.11	−0.14	0.74	−0.16	−0.58
Currency sensitivity	−0.04	0.00	−0.01	−0.01	0.38	−0.02	−0.09
Yield	−0.04	0.03	0.02	0.05	0.18	0.28	1.04
Non-est universe	0.20	−0.19	−0.05	−0.24	1.48	−0.16	−0.59
Total				5.57	5.41	1.00	3.71

Table 4.4 Characteristics of APT-estimated risk constrained portfolios (GLER Model, 2003–2011)

	Portfolio performance criteria			
	NoRC	MRC	SRC	ACWG benchmark
Geometric mean	14.16 %	13.75 %	11.08 %	4.56 %
Information ratio	0.65	0.59	0.54	
Sharpe ratio	0.53	0.49	0.41	0.16
Excess returns	9.60 %	9.19 %	6.52 %	
STD	23.20 %	24.18 %	22.71 %	17.15 %

Where NoRC = No Risk Controls GLER L = 500, MRC = Moderate Risk Controls, SRC = Strong Risk Controls, STD = Portfolio Standard Deviation

portfolio factor returns with *t*-statistics of 6.88, 2.58, and 8.85, respectively. The GLER Model portfolios have positive exposure to value and momentum and smaller stocks are purchased. The USER and GLER portfolios and the respective attribution analyses report statistically significant active returns based on specific asset selection. The active returns of the global strategy are larger than the domestic active returns see Solnik (1974, 2000) and Guerard, Rachev, and Shao (2013) and Guerard, Markowitz, and Xu (2015).

4.6 Conclusions

Investing with fundamental, expectations, and momentum variables is a good investment strategy over the long run. Stock selection models often use momentum, analysts' expectations, and fundamental data. We find support for composite modeling using these sources of data. In Chap. 1, Guerard reported that an Axioma worldwide statistical risk model estimated with all 15 factors outperformed the statistical model using only one and four factors. The APT risk model, using all PCA-based factors, is effective in creating portfolios. We find additional evidence to APT multifactor models for portfolio construction and risk control. We develop and estimate three levels of testing for stock selection and portfolio construction. The uses of multifactor risk-controlled portfolio returns allow us to reject the data mining corrections test null hypothesis. The anomalies literature can be applied in real-world portfolio construction.

Table 4.5 Axioma fundamental risk model attribution of APT Lambda = 200 portfolio returns (FSGLER Data, 1/1999–12/2011)

Total returns								
Portfolio	*NoRCGLER*							
Portfolio	0.145							
Benchmark	0.016							
Active	0.130							
Local returns								
	Return	*Risk*	*IR*	*T-Stat*	*Beg # of assets*	*End # of assets*		
Portfolio	0.145	0.231	n/a	n/a	55	123		
Benchmark	0.016	0.204	n/a	n/a	7219	10,818		
Active	0.130	0.105	1.231	4.438	7219	10,819		
Factor/specific contribution breakdown								
Factor contribution	0.063							
Specific return contribution	0.0067							
Active return	0.130							
Return decomposition								
Contributor			*Return*	*Return*	*Return*	*Risk*	*IR*	*T-Stat*
Risk free rate			0.030					
Portfolio return			0.145					
Benchmark return			0.016					
Active return			0.130			0.105	1.231	4.438
	Market timing			0.000		n/a	n/a	n/a
	Specific return			0.067		0.071	0.948	3.417
	Factor contribution			0.063		0.078	0.803	2.894
WW21AxiomaMH.Style					0.017	0.069	0.252	0.908
WW21AxiomaMH.Market					0.000	0.000	0.547	1.972
WW21AxiomaMH.Local					0.001	0.006	0.227	0.819
WW21AxiomaMH.Industry					0.018	0.023	0.803	2.896
WW21AxiomaMH.Currency					0.006	0.016	0.342	1.231
WW21AxiomaMH.Country					0.020	0.027	0.745	2.686

Contributors to active return by WW21AxiomaMH.Style						
WW21AxiomaMH.Style	*Contribution*	*Avg Wtd Exp*	*HR*	*Risk*	*IR*	*T-Stat*
WW21AxiomaMH.Medium-Term Momentum	0.049	0.360	0.705	0.020	2.456	8.854
WW21AxiomaMH.Value	0.024	0.514	0.667	0.013	1.910	6.888
WW21AxiomaMH.Liquidity	0.007	0.249	0.583	0.010	0.715	2.577
WW21AxiomaMH.Growth	0.006	0.261	0.622	0.004	1.458	5.257
WW21AxiomaMH.Size	0.005	−0.955	0.545	0.060	0.083	0.300
WW21AxiomaMH.Exchange Rate Sensitivity	−0.001	0.086	0.506	0.003	−0.405	−1.462
WW21AxiomaMH.Leverage	−0.005	0.163	0.410	0.002	−1.895	−6.834
WW21AxiomaMH.Short-Term Momentum	−0.018	0.117	0.372	0.013	−1.406	−5.069
WW21AxiomaMH.Volatility	−0.050	0.532	0.378	0.044	−1.144	−4.123
Contributors to active return by WW21AxiomaMH.Market						
WW21AxiomaMH.Market	*Contribution*	*Avg Wtd Exp*	*HR*	*Risk*	*IR*	*T-Stat*
WW21AxiomaMH.Global Market	0.000	0.000	0.494	0.000	0.547	1.972
Contributors to active return by WW21AxiomaMH.Local						
WW21AxiomaMH.Local	*Contribution*	*Avg Wtd Exp*	*HR*	*Risk*	*IR*	*T-Stat*
WW21AxiomaMH.Domestic China	0.001	0.006	0.263	0.006	0.227	0.819
Contributors to active return by WW21AxiomaMH.Industry						
WW21AxiomaMH.Industry	*Contribution*	*Avg Wtd Exp*	*HR*	*Risk*	*IR*	*T-Stat*
WW21AxiomaMH.Communications Equipment	0.006	−0.021	0.609	0.005	1.168	4.213
WW21AxiomaMH.Metals & Mining	0.003	0.031	0.596	0.005	0.591	2.130

(continued)

Table 4.5 (continued)

WW21AxiomaMH.Health Care Providers & Services	0.003	0.020	0.513	0.005	0.563	2.031
WW21AxiomaMH.Diversified Telecommunication Services	0.002	0.011	0.532	0.004	0.652	2.350
WW21AxiomaMH.Biotechnology	0.002	0.026	0.513	0.005	0.442	1.594
WW21AxiomaMH.Oil, Gas & Consumable Fuels	0.002	0.017	0.468	0.004	0.571	2.060
WW21AxiomaMH.Semiconductors & Semiconductor Equipment	0.002	−0.015	0.487	0.005	0.386	1.390
WW21AxiomaMH.Wireless Telecommunication Services	0.002	0.025	0.538	0.003	0.599	2.158
WW21AxiomaMH.Electronic Equipment, Instruments & Components	0.001	−0.002	0.538	0.001	1.063	3.831
WW21AxiomaMH.Media	0.001	−0.006	0.628	0.001	1.422	5.126
WW21AxiomaMH.Water Utilities	0.001	0.003	0.295	0.002	0.675	2.434
WW21AxiomaMH.Real Estate Investment Trusts (REITs)	0.001	0.020	0.404	0.003	0.393	1.416
WW21AxiomaMH.Thrifts & Mortgage Finance	0.001	0.004	0.506	0.001	0.971	3.502
WW21AxiomaMH.Computers & Peripherals	0.001	−0.028	0.506	0.005	0.208	0.749
WW21AxiomaMH.Chemicals	0.001	0.018	0.513	0.002	0.460	1.657
WW21AxiomaMH.Internet & Catalog Retail	0.001	0.002	0.564	0.001	0.553	1.993
WW21AxiomaMH.Transportation Infrastructure	0.001	0.019	0.526	0.002	0.256	0.923
WW21AxiomaMH.Construction Materials	0.000	0.003	0.538	0.001	0.646	2.329
WW21AxiomaMH.Consumer Finance	0.000	−0.002	0.487	0.001	0.502	1.810
WW21AxiomaMH.Insurance	0.000	0.015	0.519	0.003	0.137	0.492
WW21AxiomaMH.Containers & Packaging	0.000	0.004	0.577	0.001	0.260	0.939
WW21AxiomaMH.Commercial Banks	0.000	0.009	0.500	0.002	0.094	0.339
WW21AxiomaMH.Health Care Technology	0.000	0.004	0.218	0.001	0.199	0.719
WW21AxiomaMH.IT Services	0.000	−0.006	0.519	0.001	0.114	0.412
WW21AxiomaMH.Professional Services	0.000	0.000	0.449	0.001	0.247	0.889
WW21AxiomaMH.Construction & Engineering	0.000	−0.002	0.506	0.001	0.213	0.768
WW21AxiomaMH.Diversified Consumer Services	0.000	0.001	0.494	0.001	0.075	0.272
WW21AxiomaMH.Aerospace & Defense	0.000	−0.002	0.500	0.001	0.067	0.243

WW21AxiomaMH.Leisure Equipment & Products	0.000	0.001	0.538	0.001	0.100	0.360
WW21AxiomaMH.Household Durables	0.000	0.003	0.481	0.001	0.077	0.277
WW21AxiomaMH.Hotels, Restaurants & Leisure	0.000	0.017	0.596	0.002	0.028	0.102
WW21AxiomaMH.Building Products	0.000	0.003	0.526	0.001	0.078	0.280
WW21AxiomaMH.Multi-Utilities	0.000	−0.004	0.468	0.001	0.058	0.208
WW21AxiomaMH.Distributors	0.000	0.010	0.449	0.001	0.027	0.098
WW21AxiomaMH.Independent Power Producers & Energy Traders	0.000	0.001	0.417	0.000	0.046	0.167
WW21AxiomaMH.Textiles, Apparel & Luxury Goods	0.000	−0.002	0.532	0.001	−0.004	−0.016
WW21AxiomaMH.Air Freight & Logistics	0.000	−0.004	0.506	0.000	−0.029	−0.106
WW21AxiomaMH.Office Electronics	0.000	−0.006	0.500	0.001	−0.134	−0.481
WW21AxiomaMH.Energy Equipment & Services	0.000	−0.005	0.359	0.002	−0.071	−0.257
WW21AxiomaMH.Electrical Equipment	0.000	−0.007	0.449	0.001	−0.240	−0.864
WW21AxiomaMH.Real Estate Management & Development	0.000	0.009	0.519	0.001	−0.148	−0.535
WW21AxiomaMH.Trading Companies & Distributors	0.000	0.002	0.436	0.000	−0.370	−1.334
WW21AxiomaMH.Road & Rail	0.000	−0.008	0.462	0.001	−0.213	−0.769
WW21AxiomaMH.Life Sciences Tools & Services	0.000	0.001	0.218	0.001	−0.302	−1.090
WW21AxiomaMH.Automobiles	0.000	−0.003	0.532	0.001	−0.230	−0.830
WW21AxiomaMH.Software	0.000	−0.047	0.481	0.006	−0.047	−0.170
WW21AxiomaMH.Gas Utilities	0.000	0.000	0.436	0.001	−0.350	−1.263
WW21AxiomaMH.Food & Staples Retailing	0.000	−0.006	0.481	0.001	−0.335	−1.208
WW21AxiomaMH.Health Care Equipment & Supplies	0.000	0.005	0.481	0.001	−0.283	−1.019
WW21AxiomaMH.Internet Software & Services	0.000	0.003	0.487	0.002	−0.182	−0.658
WW21AxiomaMH.Food Products	0.000	−0.004	0.462	0.001	−0.517	−1.864
WW21AxiomaMH.Household Products	0.000	−0.023	0.538	0.003	−0.173	−0.622
WW21AxiomaMH.Multiline Retail	0.000	0.003	0.442	0.001	−0.362	−1.303
WW21AxiomaMH.Machinery	0.000	−0.007	0.404	0.001	−0.613	−2.211
WW21AxiomaMH.Auto Components	0.000	0.001	0.513	0.001	−0.615	−2.219
WW21AxiomaMH.Commercial Services & Supplies	0.000	0.019	0.500	0.001	−0.348	−1.255

(continued)

Table 4.5 (continued)

WW21AxiomaMH.Diversified Financial Services	0.000	0.012	0.468	0.001	−0.408	−1.471
WW21AxiomaMH.Airlines	−0.001	0.022	0.474	0.004	−0.130	−0.468
WW21AxiomaMH.Industrial Conglomerates	−0.001	−0.026	0.462	0.003	−0.207	−0.747
WW21AxiomaMH.Tobacco	−0.001	−0.009	0.423	0.001	−0.529	−1.908
WW21AxiomaMH.Personal Products	−0.001	−0.009	0.462	0.001	−0.700	−2.525
WW21AxiomaMH.Capital Markets	−0.001	0.004	0.449	0.002	−0.453	−1.634
WW21AxiomaMH.Marine	−0.001	0.012	0.391	0.003	−0.266	−0.959
WW21AxiomaMH.Paper & Forest Products	−0.001	0.006	0.558	0.001	−0.663	−2.389
WW21AxiomaMH.Electric Utilities	−0.001	0.007	0.449	0.002	−0.602	−2.169
WW21AxiomaMH.Pharmaceuticals	−0.001	−0.084	0.474	0.010	−0.110	−0.395
WW21AxiomaMH.Specialty Retail	−0.001	0.000	0.506	0.002	−0.602	−2.170
WW21AxiomaMH.Beverages	−0.002	−0.032	0.404	0.003	−0.563	−2.028

References

APT Analytics Guide. (2011). London: SunGard APT. www.sungard.com/apt/learnmore/

Basu, S. (1977). Investment performance of common stocks in relations to their price earnings ratios: A test of market efficiency. *Journal of Finance, 32*, 663–682.

Beaton, A. E., & Tukey, J. W. (1974). The fitting of power series, meaning polynomials, illustrated on bank-spectroscopic data. *Technometrics, 16*, 147–185.

Blin, J. M., Bender, S., & Guerard, J. B., Jr. (1997). Earnings forecasts, revisions and momentum in the estimation of efficient market-neutral Japanese and U.S. Portfolios. In A. Chen (Ed.), *Research in finance* (Vol. 15, pp. 93–114). Greenwich, CT: JAI Press.

Bloch, M., Guerard, J. B., Jr., Markowitz, H. M., Todd, P., & Xu, G.-L. (1993). A comparison of some aspects of the U.S. and Japanese equity markets. *Japan and the World Economy, 5*, 3–26.

Brown, L. D. (1999 and 2008). Annotated I/B/E/S Bibliography.

Brush, J. S. (2001). *Price momentum: A twenty-year research effort.* Columbine Newsletter Special Issue.

Brush, J. S. (2007). A flexible theory of price momentum. *Journal of Investing, 16*, 36–42.

Brush, J. S., & Boles, K. E. (1983). The predictive power in relative strength and CAPM. Journal of Portfolio Management, 9, 20–23.

Chan, L. K. C., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *Journal of Finance, 46*, 1739–1764.

Connor, G., Goldberg, L., & Korajczyk, R. A. (2010). *Portfolio risk analysis*. Princeton: Princeton University Press.

Connor, G. (1993). The three types of factor models: A comparison of their explanatory power. *Financial Analysts Journal, 51*, 42–46.

Connor, G., & Korajczyk, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics, 21*, 255–289.

Connor, G., & Korajczyk, R. A. (1993). A test for the number of factors in an approximate factor model. *Journal of Finance, 48*, 1263–1291.

Connor, G., & Korajczyk, R. A. (1995). The arbitrage pricing theory and multifactor models of asset returns, Chapter 4. In R. Jarrow, V. Maksimovic, & W. Ziemba (Eds.), *Finance, handbooks in operations research and management science* (Vol. 9). Amsterdam: North Holland.

Cragg, J. G., & Malkiel, B. G. (1968). The consensus and accuracy of some predictions of the growth of corporate earnings. *Journal of Finance, 23*, 67–84.

Dhrymes, P. J., Friend, I., Gultekin, N. B., (1984). A critical reexamination of the empirical evidence on the APT model. *Journal of Finance, 39*, 323–346.

Dhrymes, P. J., Friend, I., Gultekin, B., & Gultekin, M. (1985). New tests of the APT and their implications. *Journal of Finance, 40*, 659–674.

Dremen, D. (1979). *Contrarian investment strategy*. New York: Random House.

Dremen, D. (1998). *Contrarian investment strategies: The next generation*. New York: Simon & Schuster.

Elton, E. J., & Gruber, M. J. (1972). Earnings estimates and the accuracy of expectational data. *Management Science, 18*, B409–B424.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzman, W. N. (2007). *Modern portfolio theory and investment analysis* (7th ed.). New York: John Wiley & Sons.

Elton, E. J., Gruber, M. J., & Gultekin, M. (1981). Expectations and share prices. *Management Science, 27*, 975–987.

Fama, E. F., & French, K. R. (1992). Cross-sectional variation in expected stock returns. *Journal of Finance, 47*, 427–465.

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *Journal of Finance, 63*, 1653–1678.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*, 607–636.

Farrell, J. L., Jr. (1974). Analyzing covariance of returns to determine homogeneous stock groupings. *Journal of Business, 47*, 186–207.

Farrell, J. L., Jr. (1997). *Portfolio management: Theory and applications*. New York: McGraw-Hill.
Graham, B., & Dodd, D. (1934). *Security analysis: Principles and technique* (1st ed.). New York: McGraw-Hill Book Company.
Grinhold, R., & Kahn, R. (1999). *Active portfolio management*. New York: McGraw-Hill.
Guerard, J. B. (2010). *The handbook of portfolio construction: Contemporary applications of Markowitz techniques, Chapter 3*. New York: Springer.
Guerard, J. B. (2012). Global earnings forecast efficiency. *Research in Finance, 28*, 19–47.
Guerard, J. B., Jr., Gultekin, M., & Stone, B. K. (1997). The role of fundamental data and analysts' earnings breadth, forecasts, and revisions in the creation of efficient portfolios. In A. Chen (Ed.), *Research in finance* (Vol. 15). Stamford, CT: JAI Press.
Guerard, J. B., Jr., Gultekin, M. N., & Xu, G. (2012). Investing with momentum: The past, present, and future. *Journal of Investing, 21*, 68–80.
Guerard, J. B., Jr., Markowitz, H. M., & Xu, G. (2015). Earnings forecasting in a global stock selection model and efficient portfolio construction and management. *International Journal of Forecasting, 31*, 550–560.
Guerard, J. B., Jr., Rachev, R. T., & Shao, B. (2013). Efficient global portfolios: Big data and investment universes. *IBM Journal of Research and Development, 57*(5), 1–11.
Gunst, R. F., Webster, J. T., & Mason, R. L. (1976). A comparison of least squares and latent root regression estimators. *Technometrics, 18*, 75–83.
Haugen, R. A., & Baker, N. (1996). Communality in the determinants of expected results. *Journal of Financial Economics, 41*, 401–440.
Haugen, R., & Baker, N. (2010). Case closed. In *The handbook of portfolio construction: Contemporary applications of Markowitz techniques*. New York: Springer.
Hawawini, G., & Keim, D. B. (1995). On the predictability of common stock returns: World-wide evidence. In R. A. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in operations research and management science: Finance* (Vol. 9, pp. 497–537). Amsterdam: North-Holland.
Keane, M. P., & Runkle, D. E. (1998). Are financial analysts' forecasts of corporate profits rational? *The Journal of Political Economy, 106*, 768–805.
Korajczyk, R. A., & Sadka, R. (2004). Are momentum profits robust to trading costs? *Journal of Finance, 59*(2004), 1039–1082.
Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation and risk. *Journal of Finance, 49*, 1541–1578.
Latane, H. A. (1959). Criteria for choice among risky ventures. *Journal of Political Economy, 67*, 144–155.
Levy, H. (1999). *Introduction to investments* (2nd ed.). Cincinnati: South-Western College Publishing.
Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance, 56*, 369–385.
Lintner, J. (1965). The valuation of risk assets on the selection of risky investments in stock portfolios and capital investments. *The Review of Economics and Statistics,* 13–37.
Lintner, J. (1965b). Security prices, risk, and the maximum gain from diversification. *Journal of Finance, 30*.
Malkiel, B. (1996). *A random walk down Wall Street* (6th ed.). New York: W.W. Norton & Company.
Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.
Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly, 3*, 111–133.
Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investment* (Cowles foundation monograph, Vol. 16). New York: John Wiley & Sons.
Markowitz, H. M. (1976). Investment in the long run: New evidence for an old rule. *Journal of Finance, 31*, 1273–1286.
Markowitz, H. M. (1987). *Mean-variance analysis in portfolio choice and capital markets*. Oxford: Basil Blackwell.
Markowitz, H. M., & Xu, G. (1994). Data mining corrections. *Journal of Portfolio Management, 21*, 60–69.

Markowitz, H. M. (2013). *Risk-return analysis*. New York: McGraw-Hill.
Maronna, R. A., Martin, R. D., & Yojai, V. J. (2006). *Robust statistics: Theory and methods*. New York: Wiley.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*, 768–783.
Markowitz, H. M. (2013). *Risk-return analysis*. New York: McGraw-Hill.
Ramnath, S., Rock, S., & Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting, 24*, 34–75.
Rosenberg, B. (1974). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis, 9*, 263–274.
Rosenberg, B., & Marathe, V. (1979). Tests of capital asset pricing hypotheses. In H. Levy (Ed.), *Research in finance* (Vol. 1). Greenwich, CT: JAI Press.
Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341–360.
Ross, S. A., & Roll, R. (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance, 35*, 1071–1103.
Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. New York: John Wiley & Sons.
Rudd, A., & Clasing, H. K. (1982). *Modern portfolio theory: The principles of investment management*. Homewood, IL: Dow-Jones Irwin.
Rudd, A., & Rosenberg, B. (1980). The ‘market model’ in investment management. *Journal of Finance, 5*, 597–607.
Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science, 9*, 277–293.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business: A Supplement No. 1 Part 2,* 119–138.
Solnik, B. (1974). Why not diversify internationally rather than domestically. *Financial Analysts Journal, 38*, 48–54.
Solnik, B. (2000). *International investments* (4th ed.). Reading, MA: Addison Wesley Longman.
Stone, B. K. (1970). *Risk, return, and equilibrium: A general single-period theory of asset selection and capital market equilibrium*. Cambridge, MA: MIT Press.
Stone, B. K. (1973). A linear programming formulation of the general portfolio selection problem. *Journal of Financial and Quantitative Analysis, 8*, 621–636.
Stone, B. K. (2003). *Rank-based security grouping,multicolinearity, and the assessment of cross-sectional return dependencies.* Working Paper. The Marriott School of Management, Brigham Young University.
Stone, B. K. (2016). Portfolio performance assessment: Statistical issues and methods for improvement. In J. B. Guerard (Ed.), *Portfolio construction, measurement and efficiency*. New York: Springer.
Stone, B. K., & Guerard, J. B., Jr. (2010). Methodologies for isolating and assessing the portfolio performance potential of stock return forecast models with an illustration. In J. B. Guerard Jr. (Ed.), *The handbook of portfolio construction: Contemporary applications of Markowitz techniques*. New York: Springer.
Theil, H. (1966). *Applied economic forecasting*. Amsterdam: North-Holland.
Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.
Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market. *Harvard Business Review, 44*, 131–136.
von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
Wheeler, L. B. (1994). Changes in consensus earnings estimates and their impact on stock returns. In B. Bruce & C. B. Epstein (Eds.), *The handbook of corporate earnings analysis*. Chicago, IL: Probus Publishing.
Williams, J. B. (1938). *The theory of investment value*. Cambridge: Harvard University Press.

Wormald, L., & van Der Marwe, E. (2012). Constrained optimization for portfolio construction. *Journal of Investing, 21*, 44–59.

Yohai, V. J. (1987). High breakdown point and high efficiency robust estimates for regression. *Annals of Statistics, 15*, 642–656.

Yohai, V. J., Stahel, W. A., & Zamar, R. H. (1991). A procedure for robust estimation and inference in linear regression. In W. A. Stahel & S. W. Weisberg (Eds.), *Directions in robust statistics and diagnostics, part II*. New York: Springer.

Chapter 5
Validating Return-Generating Models

Marshall E. Blume, Mustafa N. Gültekin, and N. Bülent Gültekin

Return-generating models and the assessment of conditional expected returns underlie many important applications in finance. Jensen's (1969) measures of investment performance, as well as those based upon the Ross' (1976) Arbitrage Pricing Theory (APT), compare the realized return of a portfolio to a benchmark return. The benchmark return is an expected return conditional on some set of publicly available information. The assessment of a conditional expected return requires the specification of some stochastic process to characterize realized returns. Likewise, studies of the effect of an announcement of an unanticipated event often measure this effect by the difference between the realized return at the time of the announcement and some conditional expected return. Again, this measurement requires the specification of some stochastic process.

An assumption underlying many studies is that the market model, or more generally a model with one factor common to all securities, generates realized returns. In such a one-factor model, realized returns are the sum of an asset's response to a stochastic factor common to all assets and a factor unique to the individual asset. In the last decade, there has been much interest in models with more than one common stochastic factor, using either pre-specified factors, like Fama and French (1993) 3-factor model, or factors identified through factor analysis or similar multivariate techniques.[1]

---

[1]Factor analysis and similar factor analytic techniques have on occasion played an important role in the analysis of returns on common stocks and other types of financial assets. Farrar (1962)

M.E. Blume • N.B. Gültekin
The Wharton School, University of Pennsylvania, 3730 Walnut St., Philadelphia, PA 19104, USA

M.N. Gültekin (✉)
The Kenan-Flagler Business School, The University of North Carolina at Chapel Hill,
300 Kenan Dr, Chapel Hill, NC 27599, USA
e-mail: gultekin@unc.edu

© Springer International Publishing Switzerland 2017
J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_5

A typical way to use a return-generating model is to estimate the model with data from one period of time and then employ the estimated model to calculate conditional expected returns in a different period of time—often the immediately following period. Implicit in this use of a model is the assumption that the underlying model is stationary over time. In fact, it is highly unlikely that any economic model, except for the most trivial, is stationary over time. The question is not whether a model is stationary, but rather the degree of sensitivity to non-stationarity since the accuracy of a predictive model hinges upon the "degree" of non-stationarity.

This paper will explore the effects of such non-stationarities upon the accuracy of conditional expectations assessed for time periods following the estimation period. To this end, this paper will assess the relative accuracy of the conditional expectations of various commonly used models with data different from those used in estimating the models. In psychometrics, evaluating the accuracy of a model in terms of how it is used is termed the validation of a model.

The principal finding of this paper is that, when the criterion of accuracy is the mean-squared forecast error, multi-factor models estimated with factor analytic techniques provide more accurate out-of-sample forecasts than the Fama–French 3-factor model and the usual market model. The predictive accuracy of the market model itself depends critically on the choice of the index—equal-weighted or value-weighted. The paper also examines one model that includes the pre-specified macro variables that Chen et al. (1986) have used in a prior study. The empirical evidence indicates that a model based solely upon these macro variables provides less accurate forecasts than the usual market model. Overall, the multi-factor models provide the most accurate forecasts of those models examined.

The organization of the paper is as follows. The first section describes the design of the empirical tests and proposes the mean-squared error of the forecasts as a natural statistic to analyze in the context of performance measurement and announcement studies. The second section examines various factor models to validate the number of required factors. The third section compares the accuracy

---

may have been the first to use factor analysis in conjunction with principal component analysis to assign securities into homogeneous correlation groups. King (1966) used factor analysis to evaluate the role of market and industry factors in explaining stock returns. These two studies sparked an interest in multi-index models, and a rich body of empirical work soon emerged. Examples include Elton and Gruber (1971, 1973), Meyer (1973), Farrell (1974), and Livingston (1977), among others. The major goal of these earlier studies was to establish the smallest number of "indexes" needed to construct efficient sets.

Factor models have been used in the tests of arbitrage pricing theory and its variants. See, for example, Rosenberg (1974), Rosenberg and Marathe (1979), Roll and Ross (1980), Chen (1983), Brown and Weinstein (1983), Dhrymes et al. (1984), Dhrymes et al. (1985a,b), Gültekin and Rogalski (1985), and Cho et al. (1984), to cite a few from the large literature. A four-factor model constructed with the Dhrymes, Friend, Gültekin, and Gültekin (1985b) methodology was used in conjuction with the Bloch, Guerard, Markowitz, Todd, and Xu (1993) stock selection model to construct efficient portfolios in the U.S., See Guerard, Gültekin, and Stone (1997).

of factor analytic models to the usual market model and those using pre-specified macro variables. The final section contains concluding remarks.[2]

5.1 The Design of the Experiment

The analysis in this paper for the most part follows a two-step procedure. The first step assumes the validity of specific return-generating models and utilizes one sample of data to estimate the parameters of these models. The second step uses data in a subsequent period to validate the estimated models.

5.1.1 The Validation Criterion

There are numerous ways to validate a statistical model. The specific method of validating a model hinges upon how a researcher plans to use the model. The focus of this paper is on the use of return-generating functions in performance evaluation studies and in analysis of the reaction of stock prices to unanticipated

[2]A prior and related paper is that of Conway and Reinganum (1988). The primary purpose of their paper is to assess the adequacy of the likelihood ratio test to determine the number of factors. They use as their validation criterion the accuracy of the implied variance–covariance matrix from a factor model estimated on one sample with the estimated variance–covariance matrix from a different sample in contrast to the focus of this paper on the mean-squared forecast error of the conditional predictions. These two validation criteria are clearly related, but the one used in this paper addresses directly the way in which researchers use return-generating models in event studies and performance evaluation. The reader is referred to Chen (1988) and Stambaugh (1988) for a further discussion of the differences in these two methods of validation.

Additionally, this study shows that the number of factors and the variance–covariance matrix of returns vary substantially over time, even over the July 1962–December 1972 time period that Conway and Reinganum examine. This study explicitly adjusts summary statistics for these non-stationarities.

There is also a significant difference in the selection of the estimation and validation period between this study and that of Conway and Reinganum. For the most part, Conway and Reinganum break their sample into even and odd days, using one set of days to estimate the model and the other to validate the model. This is appropriate under their assumption that the underlying variance–covariance matrix is stationary over time. They do present one analysis using the first five years to estimate a model and the second five years to validate it. This is closer to the spirit of this paper, but it still does not parallel as closely the usage of return-generating models in studies of events and performance evaluation where the prediction period is usually much shorter than the estimation period.

Finally, a major purpose of this study is to compare factor models with estimated factors, factor models with pre-specified factors, and variants of the usual market model, which was not a goal of Conway and Reinganum.

events, frequently termed "event" or "CAR" (cumulative average residual) studies. A measure consistent with these uses is the mean-squared forecast error.[3]

In his seminal article, Jensen (1969) proposes a measure of investment performance that relies upon the validity of the Capital Asset Pricing Model (CAPM) and, with the additional assumption of a one-factor generating model, shows how to estimate this measure with a least-squares regression. Implicit in least-squares regression is the objective of minimizing squared deviations.[4] Connor and Korajczyk (1988) show that Jensen's intuition generalizes to the APT and a multi-factor model. Similarly, event or CAR studies compare realized returns to conditional predicted returns, and then test the significance of the residuals using t-tests, which again use a metric based on mean-squared errors.

5.1.2 Conditional Expectations

The first part of this section develops the formulas for assessing conditional expected returns, assuming that the return-generating process for securities is jointly normal, stationary, and independent over time and that the parameters of the joint distribution are known. The second part incorporates factor models into the formulas and interprets factor models as placing restrictions on the estimated covariance matrix.

In the formulas developed following notation is used; r_i is the return on asset i less its unconditional expectation, σ_{ii} is the variance of the return on asset i, and σ_{ij} is the covariance between the returns of asset i and asset j; there are N assets.

Under normality, the expectation of r_i conditional on the returns of the remaining $(N-1)$ assets is a linear function of these remaining returns. Specifically, if R^i is the vector of returns with the return of asset i deleted, the conditional expected return is given by

$$E[r_i|R^i] = \sum_{k\neq i} w_k r_k \ , \tag{5.1}$$

where w_k are weights appropriate to asset k. From normal theory, the weights themselves are given by

[3]Other uses would suggest different criteria. An index arbitrageur might want to use a return-generating model to construct a portfolio of a limited number of stocks to mimic the S&P 500 index. In this case, a natural way to evaluate a model is to use the estimated model to form a portfolio of securities that maximizes the correlation of its return with the S&P 500 and at the same time matches the variance of the S&P 500. One way to validate such a model is to compare in a subsequent period the profits from an arbitrage strategy using the mimicking portfolio with those using all 500 stocks.

[4]A Bayesian justification of the use of a mean-squared error rests upon an investor loss function. If an investor's loss function is quadratic, the natural measure of loss is the mean-squared error.

$$W^i = (\Sigma^i)^{-1} C^i \tag{5.2}$$

where W^i is a column vector of the $(N-1)$ weights, C^i is a column vector of the covariances of the returns of asset i with respect to each of the other $(N-1)$ assets, and Σ^i is a square matrix with dimension $(N-1)$ obtained by deleting the ith row and ith column of the full covariance matrix of all N securities.

The weights, given by (5.2), have the important property that they minimize the variance of r_i conditional on R^i.[5] This is not a surprising result since these weights are nothing more than the expected value of the estimated coefficients of a regression of r_i on the returns of the remaining $(N-1)$ assets. The essence of least-squares regression is to minimize mean-squared errors.

Thus, the process of estimating a least-squares regression can be viewed as consisting of two steps: First, estimate the covariance matrix of the dependent and independent variables. Second, use this estimated matrix to estimate the regression coefficients, which can then be used as the weights in Eq. (5.1). Viewing a regression this way helps clarify the role of factor models in forming conditional expectations.

Using a factor model to assess conditional expected returns is similar to a regression but with an important exception: Factor models place restrictions on the structure of the covariance matrix of returns, whereas the usual least-squares regression places no restrictions on this matrix. To develop these restrictions, consider the factor model:

$$r_{it} = \sum_{k=1}^{K} \lambda_{ik} f_{kt} + \eta_{it} \ , \tag{5.3}$$

where K is the number of factors, λ_{ik} is the so-called factor loading of asset i on factor k, f_{kt} is the score or value of factor k during interval t, and η_{it} is a mean-zero independent disturbance. The expected value of f_{kt} is zero, and it is scaled so that $\sigma(f_{kt})$ is 1.0. In addition, estimation of a factor model requires some assumption about the covariances between the different factors. The usual assumption, which is also made in this paper, is that $\text{Cov}(f_{kt}, f_{jt}) = 0, k \neq j$.

From (5.3), the variance of the return of asset i is

---

[5]The vector of weights W^i are those that minimize

$$E(r_i - W^i R^i)^2,$$

which can be rewritten as

$$\sigma^2(r_i) - W^{i'} C^i + W^{i'} \Sigma^i W^i.$$

Minimizing this expression with respect to W^i yields Eq. (5.2) above.

$$\sigma^2(r_{it}) = \sum_{k=1}^{K} \lambda_{ik}^2 + \sigma^2(\eta_{it}) \ , \tag{5.4}$$

and the covariance between the returns of assets i and j is

$$\text{Cov}(r_{it}, r_{jt}) = \sum_{k=1}^{K} \lambda_{ik}\lambda_{jk} \ . \tag{5.5}$$

Within the estimation period and within the class of linear estimators, estimates of the conditional expected returns for asset i that place no restrictions on the estimated covariance matrix will mathematically produce the minimum mean-squared errors. However, outside the estimation period, there is no guarantee that such an unrestricted estimate of the covariance matrix will yield the minimum mean-squared errors, or even the minimum expected mean-squared errors. If the restrictions that factor models impose on the covariance matrix are valid, it is possible that calculating conditional expected returns using a covariance matrix estimated with restrictions will yield lesser mean-squared errors in the prediction period than using an unrestricted estimate.

Non-stationarities complicate the story. Without restrictions, an estimate of the covariance matrix may "discover" non-existent relations among the returns. With restrictions, an estimate of the covariance matrix may be less prone to discover non-existent relations. In turn, it is possible that restrictions, even if not perfectly true, may improve the accuracy of conditional expectations out of the estimation period. Validating various models with different data from those used in estimating the models provides some insight into these two issues: restrictions on the estimated covariance matrix and the effect of non-stationarities.

5.2 The Experiment

The first part of this section describes the data. The second and third parts analyze the conditional expected returns, or predictions, based upon these models. The fourth part examines the impact of a January seasonal on the factor results. The fifth part decomposes the mean-squared forecast errors into the sources of the errors. The final part compares the predictions of factor models and standard market models with models that use prespecified macro variables of Chen et al. (1986).

5.2.1 Data

The empirical analyses use monthly returns of 82 sets of size-ranked portfolios of NYSE stocks constructed from the CRSP file. The first set consists of all securities in the CRSP files with complete data for the six years 1926 through 1931.

These securities were ranked by their market value as of December 1930 and then partitioned into twenty size-ranked portfolios with as close to an equal number of securities as possible. This process was repeated year by year to 2012. The sixth year in each set will be used to identify the set, so that the first set is the 1931 set and the last set is the 2013 set. The total number of securities used in the analysis starts at 361 for the 1931 set, increases to 763 for the 1949 set, and then gradually reaches 1790 for the 2012 set. In anticipation of the validation tests, the first five years of each data set will be used to estimate a model, and the sixth year will be used to validate the model.

An analysis of the basic data discloses dramatic changes in the variability of the returns of the portfolios over time. The variability is greatest in the 1930s, but even in the later years, the variability does change somewhat from one year to the next (Fig. 5.1). For most years, the smaller portfolios display greater variability in returns than the larger portfolios.[6] These changes in variability make summary measures of mean-squared errors misleading without some adjustment for these changes, and such adjustments will be made as discussed below.

For the Fama–French 3-factor model we use the data provided by Kenneth French.[7] The factors $R_m - R_f$, *SMB*, and *HML* are constructed from six size/book-to-market benchmark portfolios that do not include hold ranges and do not incur transaction costs.

1. $R_m - R_f$, the excess return on the market, is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate (from Ibbotson Associates).
2. *SMB* (Small Minus Big) is the average return on three small portfolios minus the average return on three big portfolios, $SMB = 1/3(SmallValue + SmallNeutral + SmallGrowth) - 1/3(BigValue + BigNeutral + BigGrowth)$.
3. *HML* (High Minus Low) is the average return on two value portfolios minus the average return on two growth portfolios, $HML = 1/2(SmallValue + BigValue) - 1/2(SmallGrowth + BigGrowth)$.

5.2.2 Factor Models

We use the maximum likelihood method to estimate the factor models; the usual way to assess the number of required factors is to rerun the procedure, successively increasing the number of factors until the χ^2 test for the goodness of fit developed by

[6]Interestingly, there is little change over time in the relative size of the portfolio consisting of the largest stocks, even though the market value of all of the portfolios increased almost tenfold from 1930 through 2012. In 1930, the market value of the portfolio with the largest companies is 51 % of the total market value of all twenty portfolios. By 2012, this number is 43 %.

[7]The data is publicly available from: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

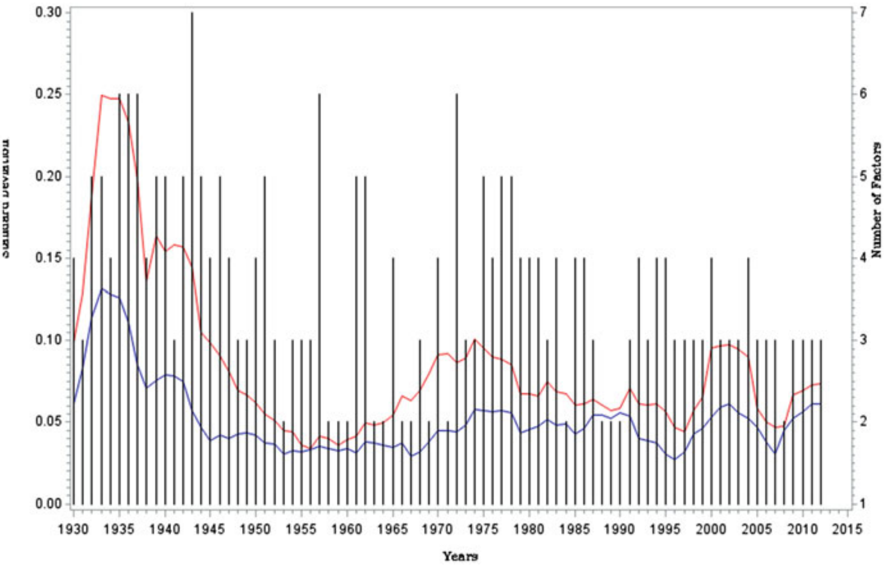


Fig. 5.1 *Portfolio return volatility and number of factors.* This figure shows the unconditional standard deviation of returns of the smallest and the largest capitalization portfolios and the adequacy of a k-factor model that generates the monthly stock returns on twenty-size ranked portfolios based on Bartlett's chi-squared test at 5X level of significance. Standard deviation and number of factors are estimated every year using the observations from the previous five years. Returns are monthly and measured as percentage changes

Bartlett (1954) indicates that the number of factors is sufficient. To use this criterion, one must specify the level of significance, often arbitrarily set at 1 or 5 %. The level of significance is important since there is a direct relation between the level of significance and the number of significant factors. However, there is no direct relation between this arbitrary level of significance and the criterion of minimizing the mean-squared errors in the forecast period.

To address the arbitrariness of setting a particular level of significance, this paper replicated the analysis for three levels of significance: 5, 10, and 20 %. For reasons to be discussed, the general nature of the results is the same whichever level of significance is used. To conserve space, the text presents only the results that use a significance level set at 5 %.

The number of required factors varies over time (Fig. 5.1). More factors are required at the beginning and the end of the 1930–2012 period than in the mid-part. Further analysis of the required number of factors reveals a positive relation between the number of factors and the variability of returns during the estimation period.[8]

[8]The Spearman's rank correlation between number of factors and the standard deviation of the equally weighted market portfolio over the sample period is 0.563, which is significant at any conventional level.

A rationale for this finding is that during periods of relatively low volatility, most of the volatility is firm-specific and it is difficult to identify the common factors. In more volatile times, the common factors are relatively more important than the firm-specific factors, making it easier to identify them.

The changing number of factors over times is strongly suggestive that the factor models are non-stationary. We conducted a series of simple Chow F-tests to formally test for stationarity.[9] We do not report these tests for brevity. The results confirm that the F-test rejects stationarity more often than could be attributed to chance and the χ^2 statistics are consistent with this impression.[10]

Since there appear to be significant non-stationarities in these factor models, validating the model with data different from those used in estimating the model is a useful tool in gaining insight into the usefulness of the model. As mentioned above, the process of validation used in this paper involves calculating conditional expectations following the estimation periods using Eqs. (5.1) and (5.2) and then analyzing the forecast errors.

The magnitudes of the mean-squared errors vary substantially over time and with portfolio size. Like the variability of the monthly returns, average mean-squared errors vary substantially for each twelve-month predictive period as a function of both time and the number of factors.[11] In view of this substantial variation, any summary measure of these mean-squared errors over time or portfolio size would be misleading without some form of scaling or normalization.

The scale factor used in this study is the mean-squared error associated with a naive forecast. The naive forecast is the average return for each portfolio in the estimation period, that is, an estimate of the unconditional expectation. An analysis of the scaled mean-squared errors shows that this normalization removes a large portion of the time trends in the annual mean-squared errors over time for a given portfolio size. However, substantial differences still remain among the size-ranked

[9]For alternative tests of non-stationarities and exploration of span of stationarity, see Hsu (1982, 1984).

[10]The stationarity test utilizes an F-statistic as proposed by Chow (1960). Although the Chow F-test was originally developed for linear regressions, it can be applied in a similar way to factor models. Specifically, estimate the factor model on the ten-year period consisting of the first five years 1926–1930 of the 1931 data set and the first five years 1931–1935 of the 1936 data set and then reestimate the factor model on the first half of the data and then on the second half, and so on. If the factor models are non-stationary, the fit of the estimated models in either half of the data will tend to be better than for the models estimated over the entire period. Specifically, if the factor model is stationary, the variances of the disturbances, η_{it} in (5.3), should be the same for all three models. The test of equality of sample variances is an F-test. Under the null hypothesis that the model is stationary and if the F-statistics are independent across portfolios, the probabilities of the F-statistics should be uniformly distributed. A χ^2-test rejects this hypothesis of an uniform distribution.

[11]For example, the average mean-squared errors for a twelve-month period ranged from 0.412 in 1944 to 26.931 in 1933 for the largest portfolio and from 1.446 in 1977 to 377.605 in 1935 for the smallest portfolio.

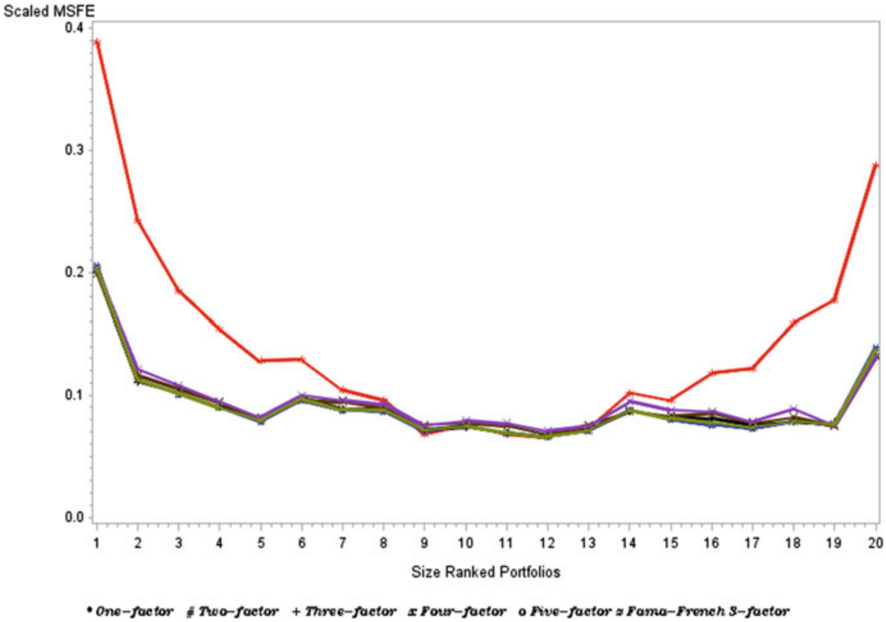


Fig. 5.2 *Scaled mean squared forecast errors for factor models* This figure shows the mean squared forecast errors scaled by the naive forecasts for size ranked portfolios. Portfolio 1 contains the smallest firms and portfolio 20 contains the largest firms. Scaled MSFE are averaged over the period 1931 to 2013

portfolios. As a consequence, the following tables and figures present summary statistics aggregated over time but not across portfolios of different sizes.

The validation of the factor models confirms the inferences based upon the χ^2 criterion that more than one factor is needed to represent the stochastic process generating returns for size-ranked portfolios. As one moves from a one-factor to a two-factor model, the mean-squared errors drop dramatically for both large and small portfolios, while there is little change for the mid-size (Fig. 5.2 and Table 5.1). As one moves to the three- or possibly four-factor model, the mean-squared errors for the large and small portfolios drop further, though only slightly. In addition, the minimum mean-squared error for the mid-size portfolios tends to occur with fewer factors than for the large or small portfolios. While we observe similar patterns for the Fama–French 3-factor model the mean-squared errors are consistently between two-factor and three-factor models.

The mean-squared errors in the forecast period for the factor models selected by the χ^2 criterion are slightly greater than the mean-squared errors associated with the best performing factor model in the forecast period for each portfolio size. The behavior of the mean-squared errors as a function of the number of factors leads to the conjecture that the arbitrary selection of two or three factors for mid-size

Table 5.1 Scaled-mean squared forecast errors for factor models for size ranked portfolios forecast periods 1931–2012

Estimation period								Forecast period								
	Factor models							Factor models						Model choice		
Portfolio size (1)	$k = 1$ (2)	$k = 2$ (3)	$k = 3$ (4)	$k = 4$ (5)	$k = 5$ (6)	$k = 6$ (7)	χ^2 at 5 % (8)	$k = 1$ (9)	$k = 2$ (10)	$k = 3$ (11)	$k = 4$ (12)	$k = 5$ (13)	FF-3 (14)	χ^2 at 5 % (15)	Min. MSFE (16)	Median k-factor (17)
1-small	0.231	0.100	0.090	0.085	0.080	0.094	0.086	0.389	0.205	0.200	0.199	0.202	0.203	0.202	0.202	4
2	0.166	0.072	0.067	0.063	0.059	0.069	0.064	0.242	0.113	0.111	0.116	0.121	0.112	0.117	0.121	3
3	0.135	0.067	0.061	0.058	0.055	0.063	0.059	0.185	0.101	0.102	0.105	0.108	0.101	0.107	0.109	3
4	0.108	0.060	0.056	0.053	0.050	0.058	0.055	0.154	0.090	0.090	0.094	0.095	0.090	0.091	0.095	3
5	0.095	0.059	0.057	0.055	0.052	0.058	0.055	0.128	0.078	0.081	0.081	0.082	0.079	0.082	0.082	3
6	0.082	0.058	0.054	0.051	0.048	0.056	0.053	0.129	0.096	0.098	0.096	0.100	0.097	0.097	0.100	3
7	0.073	0.055	0.053	0.049	0.047	0.054	0.051	0.104	0.088	0.089	0.095	0.096	0.088	0.090	0.096	3
8	0.070	0.056	0.053	0.050	0.048	0.054	0.052	0.096	0.087	0.089	0.090	0.093	0.088	0.088	0.092	3
9	0.061	0.053	0.051	0.049	0.045	0.052	0.049	0.068	0.071	0.072	0.076	0.075	0.071	0.074	0.075	2
10	0.055	0.052	0.048	0.045	0.044	0.050	0.047	0.076	0.075	0.074	0.077	0.079	0.075	0.077	0.079	2
11	0.051	0.049	0.045	0.042	0.040	0.047	0.044	0.068	0.069	0.069	0.075	0.077	0.069	0.070	0.076	2
12	0.053	0.048	0.044	0.043	0.040	0.046	0.043	0.066	0.066	0.066	0.068	0.071	0.066	0.071	0.071	2
13	0.056	0.049	0.045	0.043	0.041	0.047	0.044	0.073	0.071	0.072	0.076	0.075	0.071	0.074	0.075	2
14	0.066	0.051	0.048	0.045	0.042	0.049	0.045	0.102	0.088	0.087	0.087	0.095	0.088	0.088	0.095	3
15	0.080	0.055	0.050	0.047	0.044	0.052	0.049	0.096	0.080	0.082	0.083	0.088	0.081	0.084	0.088	3
16	0.084	0.049	0.046	0.043	0.041	0.047	0.045	0.118	0.076	0.081	0.085	0.087	0.078	0.084	0.087	2
17	0.092	0.049	0.047	0.045	0.043	0.048	0.046	0.122	0.073	0.076	0.077	0.078	0.074	0.077	0.078	2
18	0.099	0.049	0.045	0.042	0.040	0.046	0.044	0.159	0.079	0.078	0.082	0.089	0.079	0.079	0.089	3
19	0.131	0.053	0.046	0.043	0.040	0.048	0.044	0.178	0.076	0.077	0.075	0.076	0.076	0.077	0.075	3
20-large	0.202	0.088	0.077	0.071	0.065	0.080	0.074	0.288	0.139	0.131	0.131	0.132	0.137	0.136	0.130	3

(continued)

Table 5.1 (continued)

Factor model forecasts are made within and out of the sample or estimation period. Columns (2) through (9) are within sample forecast errors and Columns (8) through (16) are out of sample forecast errors. For the out of the sample period, forecasts for the month s of the year t are obtained by $r_{it+s} = \lambda_i f_{t+s}^{(i)}, s = 1, 2, \ldots, 12$, where λ_i is a $k \times 1$ vector of the factor loadings for the ith portfolio, and $f_{t+s}^{(i)}$ is a k-element column vector containing the factor scores for the prediction period $t + s$. Factor scores, $f_{t+s}^{(i)}$, are computed using the factor loadings from the estimation period and realized returns at time $t + s$. Formally, $f_{t+s}^{(i)} = \Lambda^{`(i)}[(\Lambda^{(i)}\Lambda^{`(i)} + \Omega^{(i)})]^{-1} R_{t+s}^{(i)}$, where $\Lambda^{(i)}$ is a $19 \times k$ matrix of factor loadings, $\Omega^{(i)}$ is a 19×19 diagonal matrix of specific variances, and $R_{t+s}^{(i)}$ is a 19×1 vector containing portfolio returns at time $t + s$. Superscript (i) denotes that the ith element of a vector or matrix corresponding to the ith portfolio is deleted. Returns are in deviation form from the sample mean over the estimation period. Factor model parameters, Λ and Ω are re-estimated every year using the previous 5 years' data, estimation period; forecasts and forecast errors are estimated for the next 12 months. The procedure is repeated by updating the parameter estimates of the factor models every year as a 5-year moving window. The first estimation period is from 1926 to 1930 with the corresponding forecast year being 1931. The last year of forecasts is 2013. Squared differences between realized returns and forecasts are defined as squared forecast errors. Mean squared forecast errors (MSFE) are the average of squared forecast errors for the period from 1931 to 2013

Forecast errors within sample period are the residuals of the fitted model within the same period. Within sample period forecast errors are also re-estimated every year from 1926 to 2012. Naive forecast for a year is defined as the mean portfolio returns over the estimation period

Portfolios 1 through 20 [Column (1)] contain equal number of size ranked companies. Portfolio 1 contains firms with the smallest market value while portfolio 20 the largest. Market values at the end of estimation period are used for ranking. MSFEs in Columns (2) through (7) and (9) through (14) are for one to five-factor and FF-3 models that are assumed to generate the returns. Each of these factor models is separately estimated. MSFEs in other columns are obtained from: (a) χ^2 criterion [Column (8) and (15)], a k-factor model is chosen every year based on the Bartlett's χ^2 test during the first stage of the estimation as described in Fig. 5.1; (b) Min. MSFE [Column (16)] is average of forecast errors for the factor model which produced the smallest MSFE during the estimation period. Median k-Factor model is the median number of factors that yields minimum forecast errors for out of sample period [Column (16)]; within sample period the five-factor model produces the smallest MSFEs portfolios and three or four factors for the largest and smallest portfolios leads to lesser mean-squared errors than using the standard χ^2 test.

The χ^2 criterion yields little difference in the mean-squared errors among different levels of commonly used significance, because any criterion that points to two to five factors leads to similar mean-squared errors. With a significance level of 5 %, the median number of factors over the 82 estimation periods is 4; with a significance level of 10 %, the median number is also 4; and with a significance level of 20 %, the median number is 3.

For comparison with the predictions in the forecast period, Table 5.1 also contains the average mean-squared errors for the conditional expectations within the estimation period. In contrast to the predictions in the validation period, the average mean-squared errors decrease monotonically for each portfolio as the number of factors increases from one to five. On the surface, this result suggests that the greater the number of factors the better. However, the validation of the models with additional data shows that there is little difference between models with anywhere from two to five factors.

5.2.3 The Market Model

If more than one factor in the process generates returns, the mean-squared errors from factor models should be smaller than those from the usual market model, given by

$$r_i = \beta_i r_m + \epsilon_i \tag{5.6}$$

where r_m is the return on a market index, again with all returns measured from their unconditional expectations. The measure of the market is alternatively a value-weighted or an equally weighted index of NYSE stocks. The associated covariance matrix for the market model is

$$\begin{bmatrix} \sigma^2(r_i) & \beta_i\sigma^2(r_m) \\ \beta_i\sigma^2(r_m) & \sigma^2(r_m) \end{bmatrix} \tag{5.7}$$

Applying Eq. (5.2) yields the conditional expectation $E(r_i|r_m)$ as $[\text{Cov}(r_i, r_m)/\sigma^2(r_m)]r_m$, the usual conditional forecast for the market model. It should be noted that asset i is included in the market portfolio, and thus the conditional forecast of r_i is partially conditioned by itself, a fact of importance in explaining the behavior of the mean-squared errors for the portfolio with the largest companies.

As with the factor models, there is substantial evidence of non-stationarity in the market models using either the equally weighted index or the value-weighted index. In view of this possible non-stationarity, it is appropriate to validate either variant of the market model with subsequent data. Generally, the mean-squared errors for

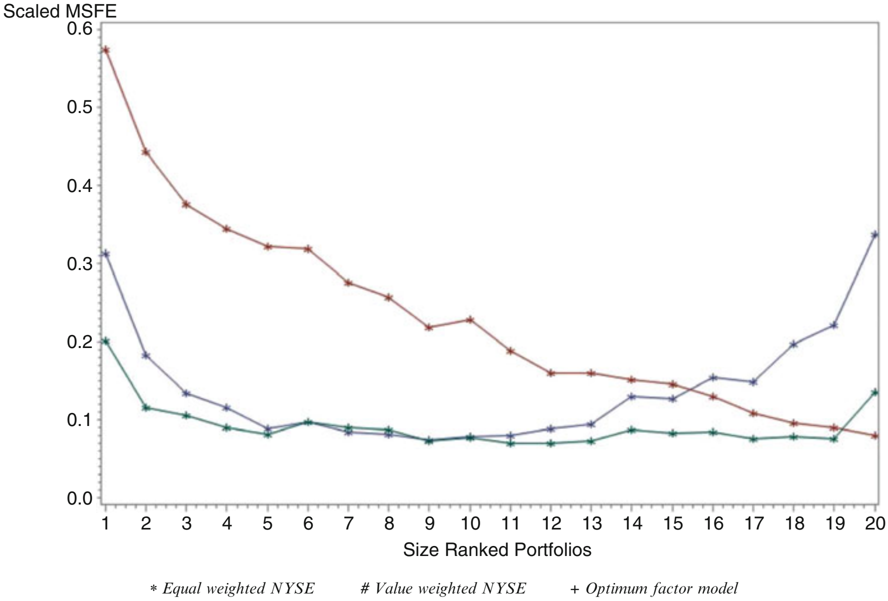


Fig. 5.3 *Scaled mean squared forecast errors for the market model and the factor model.* This figure compares the mean squared forecast errors scaled by the naive forecasts for market model and the factor model for size ranked portfolios. For market model equally and value weighted New York Stock Exchange index are used. Choice of factor model is based upon the chi–square tests of factor analysis. Scaled forecast errors are averaged over the period 1931–2012

the market model are greater than those for the factor models (Fig. 5.3). The glaring exception is the largest portfolio using a value-weighted index. Since the stocks in the largest portfolio represent an extremely large proportion of a value-weighted index of NYSE stocks and since this index is used to forecast the returns of this portfolio, this result is not surprising. Except for the largest five portfolios, the mean-squared errors associated with the equal-weighted index are less than those associated with the value-weighted index.

5.2.4 A January Seasonal

A large body of literature shows that the distribution of stock returns in January is different from the distribution of stock returns in other months. Keim (1983) found significant differences in the returns of small and large stocks in January. Tinic and West (1984) showed that virtually all of the relation between returns and betas in tests of the Capital Asset Pricing Model is due to a January seasonal. Gültekin and Gültekin (1987) demonstrated that the same is true for the two-stage tests of the Arbitrage Pricing Model.

Likewise, the factor models estimated in this paper display a January seasonal in the mean-squared errors. For every size portfolio, the mean-squared errors for January are uniformly greater than those for the other months of the year. In the case of the smallest portfolio, the mean-squared errors for January are over three times as great as the mean-squared errors for the remaining months.

This January seasonal raises the question of whether the better forecasting characteristics of a multi-factor model may be due solely to the returns in January. To answer this question, we reestimated the factor models excluding the January returns in each estimation period. According to the χ^2 criterion at a level of 5 %, the median number of required factors drops from four to three.

Even with January excluded, the minimum scaled mean-squared forecast errors still tend to occur with more than one factor (Table 5.2). For the forecast errors for February through December, two-factor models yield smaller scaled mean-squared errors than one-factor models in all cases except one mid-size portfolio. Although January was excluded in the estimating period, the estimated models still can be used to forecast January returns. For these January returns, two-factor models yield smaller scaled mean-squared errors than one-factor models in all but four cases. Thus, the presence of more than one factor is not due just to a January seasonal.[12]

5.2.5 *Biases and Inefficiencies*

Theil's decomposition shows that most of the mean-squared forecast error is random, except for the smallest portfolio, regardless of which forecasting model is used (Table 5.3).[13] The random component almost always accounts for over 90 % of the mean-squared forecast errors, and frequently accounts for over 95 % for all but the smallest portfolio.

Still, some differences among the various models warrant mention. The biases associated with the market model using an equal-weighted index of NYSE stocks are smallest for the mid-size portfolios and increase as the size of the stocks in the portfolio becomes more extreme—either larger or smaller. The largest bias, 7.6 %, is associated with the smallest portfolio. The biases for the market model using a value-weighted portfolio of NYSE stocks are similar for the mid-size portfolios

[12]To determine the importance of a January seasonal, we replicated the early analysis including, in addition to the returns on the twenty portfolios, a variable with a value of 1.0 for the months of January and 0.0 otherwise. According to the χ^2 criterion at a level of 5 %, the median number of required factors is four as before. However, there is virtually no improvement in the mean-squared errors. Again the number of factors that minimize the mean-squared error is less for the mid-size portfolios than the large or small portfolios.

Although there is a January seasonal, directly incorporating such a variable does not improve the mean-squared errors. The returns themselves already capture this seasonal. Thus, a January seasonal of itself does not account for the presence of more than one factor.

[13]See Theil (1966), and Mincer and Zarnovitz (1969) for details.

Table 5.2 Scaled mean squared forecast errors for factor models for size ranked portfolios: out of sample period forecasts adjusted for January seasonal

Forecast periods 1931–2013															
	January only					February–December					All months				
	k-factor model					k-factor model					k-factor model				
Portfolio size (1)	$k = 1$ (2)	$k = 2$ (3)	$k = 3$ (4)	$k = 4$ (5)	$k = 5$ (6)	$k = 1$ (7)	$k = 2$ (8)	$k = 3$ (9)	$k = 4$ (10)	$k = 5$ (11)	$k = 1$ (12)	$k = 2$ (13)	$k = 3$ (14)	$k = 4$ (15)	$k = 5$ (16)
1-small	1.549	0.571	0.552	0.529	0.537	0.282	0.178	0.179	0.177	0.177	0.387	0.210	0.210	0.207	0.207
2	0.871	0.219	0.212	0.199	0.200	0.186	0.104	0.105	0.110	0.114	0.243	0.114	0.114	0.117	0.121
3	0.588	0.198	0.197	0.237	0.218	0.149	0.098	0.100	0.103	0.108	0.186	0.106	0.108	0.115	0.117
4	0.500	0.135	0.120	0.113	0.123	0.125	0.084	0.085	0.088	0.092	0.157	0.088	0.088	0.090	0.095
5	0.273	0.102	0.114	0.125	0.132	0.117	0.078	0.080	0.081	0.083	0.130	0.080	0.083	0.085	0.087
6	0.309	0.136	0.136	0.129	0.137	0.114	0.093	0.095	0.094	0.100	0.130	0.096	0.099	0.096	0.103
7	0.135	0.136	0.144	0.136	0.135	0.101	0.085	0.089	0.095	0.097	0.104	0.089	0.093	0.099	0.100
8	0.133	0.102	0.098	0.100	0.119	0.093	0.085	0.088	0.090	0.092	0.096	0.087	0.089	0.091	0.094
9	0.088	0.130	0.126	0.120	0.111	0.068	0.068	0.072	0.073	0.075	0.069	0.073	0.077	0.077	0.078
10	0.101	0.107	0.111	0.107	0.112	0.075	0.073	0.072	0.077	0.079	0.077	0.076	0.075	0.080	0.082
11	0.086	0.092	0.085	0.126	0.117	0.066	0.068	0.068	0.072	0.077	0.068	0.070	0.070	0.076	0.080
12	0.145	0.118	0.116	0.122	0.105	0.059	0.062	0.065	0.067	0.071	0.066	0.067	0.069	0.072	0.073
13	0.133	0.120	0.112	0.123	0.129	0.067	0.067	0.069	0.071	0.073	0.073	0.071	0.072	0.076	0.077
14	0.217	0.150	0.147	0.126	0.124	0.093	0.087	0.088	0.089	0.096	0.103	0.092	0.093	0.092	0.098
15	0.185	0.113	0.099	0.100	0.112	0.088	0.077	0.081	0.082	0.090	0.096	0.080	0.083	0.084	0.092
16	0.380	0.138	0.146	0.165	0.164	0.093	0.071	0.075	0.080	0.081	0.117	0.077	0.081	0.087	0.088
17	0.214	0.087	0.099	0.110	0.097	0.112	0.072	0.073	0.076	0.078	0.121	0.073	0.075	0.079	0.079
18	0.498	0.117	0.120	0.122	0.112	0.134	0.079	0.081	0.088	0.089	0.164	0.083	0.084	0.091	0.091
19	0.569	0.103	0.117	0.101	0.102	0.145	0.074	0.076	0.075	0.077	0.181	0.077	0.079	0.077	0.079
20-large	1.014	0.225	0.218	0.236	0.222	0.231	0.133	0.127	0.127	0.126	0.296	0.140	0.135	0.136	0.134

Table 5.2 (Continued)

Parameters of the factor models are estimated by excluding the January return during the estimation period. Forecasts for the month s of the year t are obtained by $r_{it+s} = \lambda_i' f_{t+s}^{(i)}$, where λ_i is a $k \times 1$ vector of the factor loadings for the ith portfolio, and $f_{t+s}^{(i)}$ is a k-element column vector containing the factor scores for the prediction period $t + s$. Factor scores, $f_{t+s}^{(i)}$, are computed using the factor loadings from the estimation period and realized returns at time $t + s$. Formally, $f_{t+s}^{(i)} = \Lambda^{`(i)}[(\Lambda^{(i)}\Lambda^{`(i)} + \Omega^{(i)}]^{-1} R_{t+s}^{(i)}$, where $\Lambda^{(i)}$ is a $19 \times k$ matrix of factor loadings, $\Omega^{(i)}$ is a 19×19 diagonal matrix of specific variances, and $R_t^{(i)}$ is a 19×1 vector containing portfolio returns at time $t + s$. Superscript (i) denotes that the ith element of a vector or matrix corresponding to the ith portfolio is deleted. Returns are in deviation form from the sample mean over the estimation period. Factor model parameters, Λ and Ω are re-estimated every year using the previous 5 years' data, estimation period; forecasts and forecast errors are estimated for the next 12 months. The procedure is repeated by updating the parameter estimates of the factor models every year as a 5-year moving window. The first estimation period is from 1926 to 1930 with the corresponding forecast year being 1931. Last year of forecasts is 2013. Squared differences between realized returns and forecasts are defined as squared forecast errors. Mean squared forecast errors (MSFE) are the average of squared forecast errors for the period from 1931 to 2013. Forecast errors for the year are scaled by the root mean squared naive forecast error. Naive forecast for a year is defined as the mean portfolio returns over the estimation period

Portfolios 1 through 20 [Column (1)] contain equal number of size ranked companies. Portfolio 1 contains firms with the smallest market value while portfolio 20 the largest. Market values at the end of estimation period are used for ranking. Number of factors [Columns (2) through (16)] indicate the factor model which is assumed to generate the returns. Each of the factor models, one to five, is separately estimated

Columns (2)–(6) present mean squared forecast errors for January only, Columns (7)–(11) for 11 months from February to December, and Columns (12)–(16) for all months

Table 5.3 Decomposition of scaled mean-squared forecast errors into sources of errors forecast periods 1931–2013

	Market model				One-factor model				Five-factor model			
		Percentage source of error				Percentage source of error				Percentage source of error		
Portfolio size (1)	Scaled MSFE (2)	Bias (3)	Inefficiency (4)	Random errors (5)	Scaled MSFE (6)	Bias (7)	Inefficiency (8)	Random errors (9)	Scaled MSFE (10)	Bias (11)	Inefficiency (12)	Random errors (13)
1-small	0.313	7.64	12.05	80.32	0.389	7.26	15.94	76.80	0.202	3.19	11.15	85.66
2	0.183	4.80	5.37	89.82	0.242	5.02	9.27	85.71	0.121	0.16	3.22	96.62
3	0.135	6.25	3.54	90.21	0.185	6.61	6.91	86.48	0.108	0.79	0.81	98.40
4	0.117	3.10	2.47	94.43	0.154	4.29	5.59	90.13	0.095	0.05	0.67	99.29
5	0.089	1.48	1.04	97.48	0.128	2.51	3.26	94.23	0.082	0.02	0.63	99.34
6	0.097	0.47	2.28	97.25	0.129	1.32	4.57	94.11	0.100	0.04	1.55	98.41
7	0.085	0.00	1.52	98.47	0.104	0.33	3.50	96.18	0.096	0.29	1.13	98.59
8	0.082	0.00	0.19	99.81	0.096	0.37	1.08	98.54	0.093	0.63	0.42	98.96
9	0.075	0.20	0.05	99.75	0.068	0.11	0.47	99.42	0.075	0.15	0.05	99.81
10	0.079	0.60	2.52	96.88	0.076	0.01	3.96	96.03	0.079	0.09	2.36	97.55
11	0.080	0.90	0.92	98.18	0.068	0.08	1.43	98.49	0.077	0.04	0.88	99.08
12	0.090	4.11	0.37	95.52	0.066	2.78	0.68	96.53	0.071	1.06	1.25	97.69
13	0.094	3.47	1.43	95.10	0.073	2.08	2.05	95.88	0.075	0.34	2.02	97.64
14	0.130	3.11	0.42	96.47	0.102	2.03	0.35	97.62	0.095	0.26	0.00	99.74
15	0.128	3.75	2.85	93.40	0.096	2.77	3.33	93.90	0.088	0.24	2.45	97.31
16	0.155	2.71	1.55	95.75	0.118	1.81	1.44	96.76	0.087	0.05	0.98	98.97
17	0.149	1.94	2.44	95.62	0.122	1.04	2.22	96.74	0.078	0.24	1.29	98.47
18	0.197	3.71	0.74	95.55	0.159	2.79	0.51	96.71	0.089	0.03	0.01	99.96
19	0.222	4.77	3.04	92.19	0.178	4.09	2.67	93.24	0.076	0.22	1.86	97.93
20-large	0.337	3.02	3.14	93.84	0.288	2.34	2.61	95.06	0.132	0.32	0.45	99.23

This table shows the decomposition of scaled mean squared forecast errors (MSFE) generated by market model (using equally weighted NYSE) and one- and five-factor models. Using Theil's method, scaled MSFEs from these models (in columns 2, 6, and 10) are decomposed into three sources of errors: Bias in columns 3, 7, and 11; Inefficiency in columns 4, 8, and 12; and Random Error in columns 5, 9, and 13. Formally: $\frac{1}{N}\sum(A_t - F_t)^2 = (\overline{A} - \overline{F})^2 + (1-\beta)^2\sigma_F^2 + (1-\rho^2)\sigma_A^2$ or, MSFE = Bias + Inefficiency + Random error, where σ_A^2 is the variance for actual returns (A) and σ_F^2 is the variance for the forecasted returns (F), ρ is the correlation between actual returns and their forecasts, and β is the slope coefficient of the regression of actual returns on forecasts

MSFE are the average of squared forecast errors for the period from 1931 to 2013. Forecast errors for the year are scaled by the root mean squared naive forecast error. Naive forecast for a year is defined as the mean portfolio returns over the estimation period (not reported in the tables). However, the biases are substantially larger for the large portfolios than for the smallest portfolios.[14] The behavior of the biases for the one-factor model is similar. With two or more factors, the biases are minimal with the exception of the smallest portfolio. But even for the smallest portfolio, the percentage biases for a five-factor model (as well as a two-, three-, or four-factor model) are nearly half of those for either the market or a one-factor model.

For all the models, the forecast errors for the small portfolios display the greatest inefficiency. As the number of factors increases to two or more, the inefficiency of the forecasts declines markedly. Again, the multi-factor models' forecasting characteristics are better than either the market model or a one-factor model.

5.2.6 Macroeconomic Variables

A growing body of research uses prespecified macroeconomic variables to estimate conditional moments of stock returns. Prespecifying macroeconomic variables overcomes one of the major difficulties of factor analysis: how to associate the estimated factors with observable and economically meaningful variables. As an example, Chen et al. (1986) used some directly observable macroeconomic variables as proxies for factors in the two-stage tests of the multi-factor pricing models in a way analogous to the use of instrumental variables in regression models.

Models incorporating macroeconomic variables can be validated in much the same way as validating the market model. Estimate the model using one set of data and validate it with a different set. For each portfolio, a regression of a time series of returns on the macro variables provides the estimated model. As before, all variables are measured from their unconditional means as estimated in the estimation period, and the validation of the estimated models uses data from the 12 months following the estimation period.

Chen, Roll, and Ross provide a detailed discussion of the selection of their macro variables. Their final list of variables is the following:

1. the equal- or value-weighted NYSE index
2. the monthly growth rate of the industrial production index, measured as $\log(IPI_{t+1}/IPI_t)$, where IPI_t is the industrial production index for the month t
3. unanticipated inflation, measured as the difference between the realized inflation for the month t and the monthly T-Bill rate at the beginning of the month (see Fama and Gibbons (1984) for details)
4. the change in the term structure, measured by the difference between the return of a portfolio of long-term government bonds and the T-Bill rate

[14]The percentage biases are 9.6, 15.6, and 16.3 % for portfolios 18, 19, and 20, respectively, and 3.0 % for the smallest portfolio.

5. changing risk premia, measured by the return on BAA rated non-convertible corporate bonds less the return on a portfolio of long-term government bonds

The validation tests in this section use these same variables. Since the industrial production data are available only after 1946 and the estimation period requires five years of data, the first forecast year is 1952 and the last 2013. For comparison purposes, some of the earlier analyses have been replicated for these years.[15]

The validation process suggests that the macro variables by themselves have no forecasting power (Table 5.4), with the scaled mean-squared errors in the validation period ranging from 1.134 to 1.271. Since these statistics have been scaled by the mean-squared errors of the naive forecasts, a statistic greater than one indicates that the naive forecasts are more accurate than those using just the macro variables. Within the estimation period, the macro variables by themselves do have some explanatory power, with the scaled mean-squared errors ranging from 0.577 to 0.909. These two results imply that the regression in the estimation period found a relation that was not there, or that any relation in the estimation was not sufficiently stationary to provide forecasting power, or some combination of the two.

Adding either the equal-weighted or value-weighted index of NYSE stocks to the macro variables leads to a substantial reduction in the scaled mean-squared errors for every portfolio. As an example, the average scaled mean-squared errors for the largest portfolio is 1.160 with just the macro variables, but drops to 0.461 with the addition of the equal-weighted index. Even more accurate are the forecasts that drop the macro variables and include just a stock market index, suggesting that the macro variables merely add noise to the forecasts. Again, the multi-factor models generally yield smaller scaled mean-squared errors than either version of the market model.

5.3 Conclusions

The goal of this paper was to validate various stochastic return-generating models on data different from those used in estimating the models. The specific models analyzed were factor models, the traditional market model, and models incorporating prespecified macroeconomic variables. The principal conclusion of this paper is that factor models with two to five factors yield more accurate predictions than either the traditional market model or a one-factor model.

A model that included the prespecified macroeconomic variables used by Chen et al. (1986) had no predictive power. Thus, at least for the macro economic variables considered here, there is no gain to adding these variables to the traditional market model. But importantly, the predictions of a multi-factor model were more accurate than the market model.

Acknowledgements We thank Craig MacKinlay and Jennifer Conrad for their careful and thoughtful comments.

[15]The stationarity tests again tend to reject stationarity.

Table 5.4 Comparison of scaled mean-squared forecast errors using macroeconomic variables, market model and factor model

Forecast periods 1952–2013						
Portfolio size (1)	Macro variables (2)	Macro & equal-weighted index (3)	Equal-weighted index (4)	Macro & value-weighted index (5)	Value weighted index (6)	Optimal factor model (7)
1-small	1.247	0.386	0.334	0.759	0.625	0.204
2	1.222	0.218	0.203	0.590	0.489	0.124
3	1.227	0.175	0.157	0.520	0.425	0.108
4	1.195	0.150	0.132	0.489	0.394	0.096
5	1.202	0.117	0.104	0.443	0.370	0.091
6	1.226	0.132	0.108	0.469	0.368	0.100
7	1.271	0.132	0.096	0.460	0.328	0.103
8	1.213	0.129	0.100	0.389	0.318	0.110
9	1.215	0.093	0.069	0.335	0.253	0.075
10	1.198	0.099	0.085	0.340	0.276	0.085
11	1.181	0.098	0.083	0.285	0.229	0.079
12	1.212	0.101	0.081	0.248	0.175	0.069
13	1.214	0.098	0.088	0.214	0.165	0.070
14	1.187	0.131	0.119	0.220	0.167	0.091
15	1.134	0.158	0.132	0.187	0.155	0.088
16	1.175	0.171	0.148	0.182	0.137	0.079
17	1.140	0.191	0.158	0.144	0.120	0.081
18	1.172	0.251	0.226	0.146	0.115	0.087
19	1.160	0.282	0.248	0.118	0.101	0.083
20-large	1.160	0.461	0.391	0.116	0.097	0.168

This table compares MSFE produced by four forecasting models. The first model uses a set of macro economic variables to make conditional forecasts (column 2). The second model includes market return as an additional exogenous variable to macro economic variables (column 3 for equally weighted NYSE index and column 5 for value weighted NYSE index). The third model forecasts are conditional on market returns (column 4 for equal weighted NYSE index and column 6 for value weighted NYSE index). The fourth model uses the optimal factor model based on the Bartlett's chi-square test during estimation period (column 7)
Model Parameters are re-estimated every year using the previous 5 years' data, estimation period; forecasts and forecast errors are estimated for the next 12 months. The procedure is repeated by updating the parameter estimates of the models every year as a 5-year moving window. The first estimation period is from 1947 to 1951 with the corresponding forecast year being 1952. Last year of forecasts is 2013. Forecast errors for the year are scaled by the root mean squared naive forecast error. Naive forecast for a year is defined as the mean portfolio returns over the estimation period. Mean squared forecast errors are the average of squared forecast errors for the period from 1952 to 2013
The macro variable are

1. Unanticipated inflation.
2. Monthly growth rate of industrial production.
3. Yield differential between BAA rated corporate bonds and long-term government bonds.
4. Yield differential between long-term government bonds and T-Bills.

The indexes are for all NYSE stocks

References

Bartlett. M. S. (1954). A note on the multiplying factors for various chi-squared approximations. *Journal of the Royal Statistical Society, 16*, 296–298.

Bloch, J. B. G., Jr., Markowitz, H. M., Todd, P., & Xu, G. L. (1993). A comparison of some aspects of the U.S. and Japanese equity markets. *Japan & the World Economy, 5*, 3–26.

Brown, S. J., & Weinstein, M. (June 1983). A new approach to testing asset pricing theories: the bilinear paradigm. *Journal of Finance, 38*, 711–743.

Chen, N. -f. (December 1983). Some empirical tests of the theory of arbitrage pricing. *Journal of Finance, 38*, 1392–1414.

Chen, N. -f. (January 1988). Stable factors in security returns: identification using cross-validation - comment. *Journal of Business & Economic Statistics, 6* , 16.

Chen, N. -f., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business, 59*, 383–403.

Cho, D. C., Elton, E. J., & Gruber, M. J. (March 1984). 'On the robustness of the Roll and Ross APT methodology. *Journal of Financial and Quantitative Analysis, 19* , 1–10.

Chow, G. C. (July 1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28* , 591–605.

Connor, G., & Korajczyk, R. A. (September 1988). Risk and return in an equilibrium APT: application of a new test methodology. *Journal of Financial Economics, 21* , 255–289.

Conway, D. A., & Reinganum, M. R. (January 1988). Stable factors in security returns: identification using cross-validation. *Journal of Business & Economic Statistics, 6* , 1–15.

Dhrymes, P. J., Friend, I., & Gültekin, N. B. (June 1984). A critical reexamination of the empirical evidence on the arbitrage pricing teory. *Journal of Finance, 39* , 323–346.

Dhrymes, P. J., Friend, I., Gültekin, N. B., & Gültekin, M. N. (March 1985a). An empirical examination of the implications of arbitrage theory. *Journal of Banking and Finance, 9* , 73–99.

Dhrymes, P. J., Friend, I., Gültekin, N. B., & Gültekin, M. N. (July 1985b). New tests of the APT and their implications. *Journal of Finance, 40*, 659–675.

Elton, E. J., & Gruber, M. J. (October 1971). Improved forecasting through the design of homogeneous groups. *Journal of Business, 44*, 432–450.

Elton, E. J., & Gruber, M. J. (December 1973). Estimating the dependence structure of share prices–implications for portfolio selection. *Journal of Finance, 8*, 1203–1232.

Fama, E. F., & Gibbons, M. (1984). A comparison of inflation forecasts. *Journal of Monetary Economics, 13*, 327–348.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Farrar, D. E. (1962). *The investment decision under uncertainty*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Farrell, J. (April 1974). Analyzing covariation of returns to determine homogeneous stock groupings. *Journal of Business, 47*, 186–207.

Guerard, J. B. Jr., Gultekin, M., & Stone, B. K. (1997). The role of fundamental data and analysts' earnings breadth, forecasts, and revisions in the creation of efficient portfolios. In: Chen, A. (Ed.), *Research in finance*, vol. 15, pp. 69–92.

Gültekin, M. N., & Gültekin, N. B. (December 1987). Stock return anomalies and the tests of the arbitrage pricing theory. *Journal of Finance, 42*, 1213–1224.

Gültekin, N. B., & Rogalski, R. J. (March 1985). Government bond returns, measurement of interest rate risk, and the arbitrage pricing theory. *Journal of Finance, 40*, 43–61.

Hsu, D. A. (March 1982). A Bayesian robust detection of shift in the risk structure of stock market returns. *Journal of the American Statistical Association, 77*, 29–39.

Hsu, D. A. (March 1984). The behavior of stock: Is it stationary or evolutionary? *Journal of Financial and Quantitative Analysis, 19*, 11–28.

Jensen, M. (April 1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business, 42*, 167–247.

Keim, D. B. (1983). Size related anomalies and stock return seasonality: further empirical evidence. *Journal of Financial Economics, 12*, 13–32.
King, B. F. (January 1966). Market and industry factors in stock price behavior. *Journal of Business, 39*, 139–190.
Livingston, M. (June 1977). Industry movements of common stocks. *Journal of Finance, 32*, 861–874.
Meyer, S. (June 1973). A re-examination of market and industry factors in stock price behavior. *Journal of Finance, 8*, 645–709.
Mincer, J., & Zarnovitz, V. (1969). The valuation of economic forecasts. In J. Mincer (Ed.), *Economic forecasts and expectations*. New York: National Bureau of Economic Research.
Roll, R., & Ross, S. A. (December 1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance, 35*, 1073–1103.
Ross, S. A. (December 1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341–360.
Rosenberg, B. (1974). Extra-Market Components of Covariance in Security Returns. *Journal of Financial and Quantitative Analysis, 9,* 263–274.
Rosenberg, S. A., & Marathe, V. (1979). Tests of capital asset pricing hypotheses. In H. Levy (Ed.), *Research in finance* (Vol. 1). Greenwich: JAI Press.
Stambaugh, R. F. (January 1988). Stable factors in security returns: identification using cross-validation - comment. *Journal of Business & Economic Statistics, 6*, 20–21.
Theil, H. (1966). *Applied economic forecasting*. Amsterdam: North-Holland Publishing Company.
Tinic, S. M., & West, R. R. (1984). Risk and return: January vs. the rest of the year. *Journal of Financial Economics, 13*, 561–574.

Chapter 6
Invisible Costs and Profitability

Xiaoxia Lou and Ronnie Sadka

6.1 Introduction

Ever since the signing of the Buttonwood Agreement on May 17, 1792, which marked the beginning of what became today's New York Stock Exchange (NYSE), investors have been constantly seeking to develop new and interesting ways to profit from various investment strategies. Libraries are filled with textbooks discussing various investment approaches; a notable example is the influential manuscript of Graham and Dodd (1934).

The academic literature in this area has long posited seemingly conflicting puzzles. The efficient market hypothesis, proposed in Fama (1970), asserts that markets are informationally efficient and, therefore, investors cannot earn abnormal returns. Grossman and Stiglitz (1980) further introduce the concept of near efficiency, arguing that because information is costly, prices only partially reflect information. Therefore, arbitragers can earn extra returns by expending resources to gather information. The academic literature has documented many trading strategies over the past several decades that generate positive risk-adjusted returns (alphas) (see McLean and Pontiff 2014, for a recent summary of dozens of such strategies and their performance after publication in the public domain). At the same time, the literature documents that most active investors have not been able to outperform passive investment strategies, such as investing in the market portfolio (see Fama and French (2010) for a recent review and new evidence).

X. Lou
Lerner College of Business and Economics, University of Delaware, Newark, DE 19716, USA
e-mail: lous@udel.edu

R. Sadka (✉)
Department of Finance, Carroll School of Management, Boston College,
140 Commonwealth Ave., Chestnut Hill, MA 02467, USA
e-mail: sadka@bc.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_6

Treynor offers his insights on these matters in a series of articles, which generally argue that transaction costs can explain why portfolio strategies consistently generate remarkable performances on paper, while their actual returns fail to beat market averages. He discusses the roles of different market participants and their interaction resulting in transaction costs. For example, Treynor (1971) points out that any transaction with a market maker incurs a bid/ask spread cost in addition to any explicit brokerage commission. The market maker interacts with three types of traders: information-based traders, liquidity-motivated traders, and noise traders (those who believe they have new information, while, in fact the information has already been incorporated into the stock price). The market maker loses to information-based traders and gains from liquidity-motivated and noise traders. To justify market making, the bid/ask spread must be large enough to cover the losses incurred while trading with informed traders. On average, therefore, the uninformed trader loses in the trading game. The bid/ask spread appears to be *hidden* "because oscillations between 'bid' and 'asked' are camouflaged by the constant fluctuations in the equilibrium value of the stock."

Treynor (1981) points out that even without a designated market maker, traders still incur transactions costs. He points out that trading is a zero-sum game, and that most active investors have not been successful because they have not excelled at playing the game. To succeed in the trading game, Treynor argues that the active investor needs to know that the trader on the other end of the trade is not drawn from a random sample. Instead, this trader has one of two motives: information or value. Information traders trade based on information, which affects the fundamental value of a stock, whereas the latter type of trading is motivated by value traders' perceived discrepancy between the market price and intrinsic value. Information-based traders are time sensitive and demand quick execution, before their possessed information is impounded into stock prices. The value traders, therefore, essentially act as liquidity providers, and the buy-sell spread can be viewed as compensation for providing liquidity. In equilibrium, the transaction costs paid by informed traders will be equal to their initial information advantage.

In Treynor (1987), both dealers and value-based traders act as market makers. The bid/ask spread of value-based traders (the outside spread) is higher than that of dealers (the inside spread). Dealers are valuable because they reduce spreads and improve the liquidity of the markets; however, dealers are limited in the amount of capital and the ability to absorb losses, thus setting limits on the positions they are willing to take to limit the risk of, in Treynor's words, being "bagged." When their limits are reached, dealers will lay off to the value-based traders, who act as market makers of the last resort. The inside spread is a function of the lay-off price that dealers face while transacting with value-based traders and is therefore tied to the outside spread.

Treynor makes a more specific argument about performance and transactions costs in Treynor (1994). He clarifies that the great majority of trading has adversarial motives. A trade typically involves the exchange of time for price, where time means the right to transact quickly. Value traders sell time and information traders buy time. Both types of traders (the information trader and the value trader) face risk.

To succeed in the trading game, one needs to know both the value and the cost of time. In practice, value (information) traders often only know the cost (value) of time, not the value (cost) of time. The value trader can, therefore, get bagged and the information trader can pay an exorbitant price for trading quickly. Treynor coined these *invisible* trading costs, and argued that they are higher than visible costs, such as commissions and market impact. Value traders who lose consistently are those who sell time for less than it is worth. Similarly, information traders who consistently lose are those who buy time for more than it is worth.

In general, Treynor stresses that to succeed in the trading game, value (information) investors must learn the value (cost) of time without revealing the cost (value) of time in the strategic trading game. In this regard, Kyle (1985) posits that informed traders partially reveal the value of time through the size of their trades. Therefore, the non-proportional price impact, that is Kyle's Lambda, can be used to gauge the invisible costs discussed by Treynor. The notion of invisible costs of trading has been the focus of numerous papers providing theoretical explanations[1] and empirical estimates[2] of such costs. In the early years of the previous decade, a growing literature on the limits of arbitrage opportunities examined whether strategies constructed to exploit several prominent asset-pricing anomalies can still remain profitable after taking into account transaction costs. For example, Knez and Ready (1996) find that size-based strategies are too costly to trade; Mitchell and Pulvino (2001) analyze the profits to risk arbitrage of mergers and acquisitions; and Korajczyk and Sadka (2004) examine the profitability of momentum strategies.[3] These empirical studies argue that transaction costs can impose a first-order effect on profitability.

The purpose of this paper is to provide synthesized thoughts about the profitability of various asset-pricing anomalies under transactions costs. Section 6.2 introduces different trading cost measures. Section 6.3 discusses performance net of transactions costs. The profitability of some asset-pricing anomalies under transactions costs is reviewed in Sect. 6.4. Section 6.5 includes evidence on the time series of liquidity and institutional investment horizon, and Sect. 6.6 concludes.

6.2 Measures of Trading Cost

There are several components of trading costs that differ dramatically both in size and ease of measurement. The components that can be measured with the least error are the explicit trading costs of commissions and bid/ask spreads. When

[1]See, e.g., Kyle (1985), Glosten and Milgrom (1985), Easley and O'hara (1987), Admati and Pfleiderer (1988), and Huberman and Stanzl (2004).

[2]See, e.g., Glosten and Harris (1988), Hasbrouck (1991a,b), Breen et al. (2002), and Sadka (2006).

[3]See also Chen et al. (2005), Lesmond et al. (2004), Mendenhall (2004), Sadka and Scherbina (2007), Ng et al. (2008), and Chordia et al. (2009).

trading an institutional-size portfolio, however, these proportional costs can be modest compared to both the additional nonproportional cost of price impact and the invisible cost of post-trade adverse price movement (Treynor 1994).

6.2.1 Proportional Costs

The two main measures of proportional cost are quoted spread (QS) and effective spread (ES). The quoted spread is the difference between the bid and ask prices and it measures the cost of a round-trip transaction, assuming trades are executed at the quoted prices. The cost of a single trade is often estimated by the quoted percentage half-spread (*QHS*), half the difference between the best ask and bid prices divided by the average of the best ask and bid prices:

$$QHS = \frac{Q^A_{i,t} - Q^B_{i,t}}{2M_{i,t}},$$

where $Q^B_{i,t}$ and $Q^A_{i,t}$ are the bid and ask price for security i at time t, respectively, and $M_{i,t}$ is the bid/ask midpoint.

The effective (half) spread is defined as the difference between the actual transaction price and the bid/ask midpoint. The effective percentage half-spread (*EHS*) is the absolute value of the difference between the transaction price and the midpoint of quoted bid and ask prices, divided by the bid/ask midpoint:

$$EHS = \frac{d_k \cdot (P_{i,k} - M_{i,t})}{M_{i,t}}, \tag{6.1}$$

where d_k is an indicator variable that equals one if trade k is buyer-initiated and negative one if seller-initiated, and $P_{i,k}$ is the transaction price of the trade.

Effective spreads may differ from quoted spreads. Transactions smaller than the quoted depth, often executed within the bid and ask prices and, hence, exhibit "price improvement" (Fialkowski and Petersen 1994). Conversely, trades larger than the quoted depth generally execute outside of the bid/ask spread (Knez and Ready 1996). Empirically, the effective spread is smaller than the quoted spread, on average.

Effective spread can be further decomposed into a realized component and a price impact component. The price impact component is defined as:

$$PI = \frac{d_k \cdot (M_{i,t+n} - M_{i,t})}{M_{i,t}},$$

where $M_{i,t+n}$ is the bid/ask midpoint n seconds after the trade time t. It measures the permanent price movement following the trade within a given look-ahead window n. The literature has used 1, 5, and 30 min for the length of look-ahead window. This is the component of transaction costs that compensates the liquidity provider for the adverse selection problem. The difference between effect spread and price impact, the realized spread, is then the compensation f or liquidity providers for the non-informational cost and risk such as the order processing cost and the inventory risk.

6.2.2 Nonproportional Costs

The economic importance of price impact is demonstrated by Loeb (1983), Keim and Madhavan (1996, 1997), and Knez and Ready (1996), who show that transaction costs increase substantially as the size of an order increases. Motivated by the linear pricing rule of Kyle (1985) which states that price change is a linear function of net volume, price impact is often estimated by the slope coefficient λ in regressions of the following form:

$$\Delta p_{i,t} = \lambda_i \cdot d_{i,t} \cdot \text{Trade Size}_{i,t} + \varepsilon_{i,t}, \tag{6.2}$$

where i indexes stock; the dependent variable $\Delta p_{i,t}$ can be percentage return, change in price, or log return; $d_{i,t}$ measures trade direction, an indicator variable that equals one if the trade is buyer-initiated and negative one if seller-initiated; t can be a calendar time interval such as a 5-min interval or in some models is defined in terms of trades.

In Hasbrouck (2009), t denotes a 5-min time interval, the dependent variable Δp_t is the log change in the quote midpoint over a 5-min period t, and $SVol_t$ is the sum of the signed square-root of dollar volume in the 5-min period:

$$\ln(m_t) - \ln(m_{t-1}) = \lambda_i^H \times \sum_k d(v_k)\sqrt{|v_k|} + \varepsilon_{i,t}, \tag{6.3}$$

where m_t is the quote midpoint at the end of the time period t; where k denotes the k^{th} trade in the interval t; $d(v_k)$ is an indicator variable that equals one if the trade is buyer-initiated and negative one if seller-initiated; v_k is the dollar volume of k^{th} trade in the 5-min interval; and λ_i^H is the asset i's price impact coefficient.

In Breen et al. (2002), the time interval used is 30 min, the dependent variable is the percentage return, and trade size is measured by net share turnover over the 30-min time periods:

$$\frac{\Delta p_{i,t}}{p_{i,t-1}} = \lambda_i^{BHK} \times Turnover_{i,t}, \tag{6.4}$$

The model implies a convex price impact function (Korajczyk and Sadka 2004).

The Glosten and Harris (GH) specification allows a decomposition of the price impact into fixed and variable components. One way to motivate the model is to assume that fundamental value at trade t, $m_{i,t}$, is equal to fundamental value at trade $t-1$ plus two types of news. One piece of news is the revision in perceived fundamental value due to the information that trade t occurred for $q_{i,t}$ shares (with the sign of $q_{i,t}$ denoting the direction of the trade initiator). This is given by the price impact coefficient times $q_{i,t}$. The other piece of news is information released that is unrelated to the trade at t, $\varepsilon_{i,t}$.

$$m_{i,t} = m_{i,t-1} + \lambda_i^{GH} q_{i,t} + \varepsilon_{i,t} \tag{6.5}$$

However, the actual trade price differs from the fundamental value by a temporary component that compensates those making a market in the stock for the non-information related costs of market making. The sign of a trade is denoted $d_{i,t}$ and is assigned a value of $+1$ for a buy and -1 for a sell. The actual traded price, $p_{i,t}$, differs from $m_{i,t}$ by $d_{i,t}\Psi_i$:

$$p_{i,t} = m_{i,t} + d_{i,t}\Psi_i. \tag{6.6}$$

In other words, the traded price is higher than the fundamental value by Ψ_i for buyer-initiated trades and is lower than the fundamental value by Ψ_i for seller-initiated trades. In essence, Ψ_i is the effective half-spread for very small trades. The observed price change from trade $t-1$ to trade t, given by:

$$\Delta p_{i,t} = \lambda_i^{GH} q_{i,t} + \Psi_i \Delta d_{i,t} + \varepsilon_{i,t}, \tag{6.7}$$

where $\Delta p_{i,t}$ is the price change of stock i from trade $t-1$ to trade t as a consequence of a (signed) trade of $q_{i,t}$ shares of the stock and the difference between the sign of a current trade and the previous trade is denoted $\Delta d_{i,t}$. The regression coefficient λ_i^{GH} represents the variable cost of trading, which has a permanent effect on price. Similarly, Ψ_i represents the fixed cost, which has a temporary effect on the price. The effective half-spread for small trades is Ψ_i and the price impact of large trades is measured by λ_i^{GH}.

6.2.3 Estimation Issues

The estimation of transaction costs has a few challenges. For example, the trade direction is typically not observed, except for a few proprietary datasets. The most common trade classification method is the (Lee and Ready 1991) algorithm which classifies a trade as a buyer-initiated trade if the price is above the quote midpoint, a seller-initiated trade if the price is below the quote midpoint, and uses the tick test if the price is equal to the midpoint. The other two methods are from Ellis et al. (2000) and Chakrabarty et al. (2007).

Many measures of transactions costs require the use of intraday data. The ISSM (Institute for the Study of Security Markets) and TAQ (Trade and Quote) datasets provide tick-by-tick data, but only starting in 1983 (ISSM: 1983–1992, and

TAQ thereafter). In addition, the size of the TAQ data has grown exponentially. Researchers have developed models and methods to estimate both proportional and nonproportional transaction costs using the daily CRSP data, which is available for a much longer time period and has lower demand on computing power. For example, Hasbrouck (2009) proposes a Bayesian approach to estimate the effective bid-ask spread using the Center for Research in Security Prices (CRSP) daily data. Corwin and Schultz (2012) propose a spread proxy based on daily high-low ratio. The Amihud (2002) measure is built upon the daily absolute return-to-dollar volume ratio to capture price impact. Goyenko et al. (2009) run a horse race between various proxies from the CRSP daily data and several high-frequency benchmarks calculated using the TAQ data and Rule 605 data. They find that several daily proxies can provide fairly good estimates for proportional transaction costs such as the percent effective and realized spreads. However, nonproportional price impact measures using daily data fail to capture the level of the high-frequency benchmarks for price impact, although the Amihud measure (2002) has reasonably high correlations with the high-frequency price impact measures.

6.3 Measures of Performance Under Transaction Costs

Transaction costs result in a reduction of the net returns earned by a trading strategy. For proportional transactions cost models, a trading strategy's performance is independent of the size of the portfolio; however, for nonproportional price impact transactions costs, such costs, as a percentage of trade size, grow with the size of the portfolio being traded; therefore, the performance of the trading strategy declines with the size of the portfolio.

Studies that examine the impact of nonproportional transaction costs on trading profitability are often interested in determining the amount that a single portfolio manager could invest before the performance of trading strategies breaks even with that of the benchmark. For example, in Korajczyk and Sadka (2004), the benchmark asset returns are the Fama-French three-factors; they calculate the size of the portfolio that (1) eliminates the statistical significance of the portfolio abnormal return, (2) drives the level of abnormal return to zero, and (3) drives the portfolio Sharpe ratio (Sharpe 1966) to that of the maximal Sharpe ratio obtained from combinations of the Fama and French (1993) three-factor portfolios. Novy-Marx and Velikov (2014) point out that the use of the common notion of abnormal return relative to a set of explanatory asset returns such as the Fama-French factors can be somewhat problematic because the factors used to calculate traditional alphas are themselves gross returns, without considering transaction costs. They therefore propose a generalized definition of alpha to better evaluate the performance of a strategy.

6.4 Are Return Anomalies Robust to Trading Cost?

6.4.1 Return Anomalies

The academic literature has documented many strategies over the past several decades that generate positive risk-adjusted returns (alphas). Many of these strategies are based on cross-sectional relations between some predetermined variables and future stock returns. The long list of predictors includes measures such as market capitalization, price-to-fundamental ratios, firm profitability, net share issuance, prior returns, idiosyncratic volatility, total accruals, asset growth, and investment-to-assets. McLean and Pontiff (2014) find that the academic literature has documented at least 95 such cross-sectional predictors in 78 different papers.

The most notable anomalies are the size premium, the value premium, momentum, and the post-earnings-announcement-drift. The size premium refers to the historical tendency of stocks with low market capitalizations to outperform stocks with high market capitalizations (Banz 1981). The value premium refers to phenomenon that stocks with low price to fundamental ratios (value stocks) outperform stocks with high price-to-fundamental ratios (growth stocks) (Fama and French 1992; Stattman 1980). The momentum, or relative strength, literature finds that, over horizons such as a week or a month, stock returns have negative serial correlation (reversals), while at three to twelve month horizons, they exhibit positive serial correlation (momentum). Over longer three-year to five-year horizons stock returns again exhibit reversals. The momentum of individual stocks is extensively examined by Jegadeesh and Titman (1993, 2001) and Chan et al. (1996, 1999). They show that one can obtain superior returns by holding a zero-cost portfolio that consists of long positions in stocks that have outperformed in the past (*winners*), and short positions in stocks that have underperformed during the same period (*losers*). Asness et al. (2013) document value and momentum phenomena across several markets and asset classes. The post-earnings-announcement drift (PEAD), or earnings momentum, was first documented by Ball and Brown (1968). Stocks that just announced surprisingly good earnings subsequently outperform those that just announced surprisingly bad earnings, that is, stock prices continue to drift in the direction of earnings surprises for several months after the announcement dates.

The literature has proposed risk-based explanations for many of the aforementioned anomalies. For example, Sadka (2006) argues that momentum is partially explained by a liquidity risk premium; however, risk-based models have not been able to completely explain away many of these anomalies including momentum. Indeed, Fama and French (1996) view this unexplained persistence of intermediate-term momentum returns throughout the last several decades as one of the most serious challenges to the notion of market efficiency.

In the absence of risk premia-based explanations for cross-sectional return anomalies, an important question is whether there are significant transaction costs that prevent investors from trading sufficiently large quantities to drive away the apparent profits. While transaction costs do not explain the underlying causes for the existence of seemingly profitable strategies, they may be sufficient to explain their persistence.

The literature has also found that most of these anomalies are more significant among small and illiquid stocks. For example, Bhushan (1994) documented that the PEAD is stronger for smaller, low-priced stocks. Jegadeesh and Titman (2001) find that the zero-cost portfolio that is long in the top momentum portfolio and short in the bottom momentum portfolio generates an average return of 1.42 % per month among small firms in 1965–1998, compared to 0.86 % among large firms.

6.4.2 Performance Net of Transaction Costs

Incorporating explicit trading costs, such as commissions and spreads, into portfolio returns has occurred in the literature for some time. Incorporating nonproportional price impacts into trading strategies has only recently received significant attention. After intraday trade and quote data became available, many studies now rely on TAQ data to obtain estimates of transaction costs, while several papers, such as Frazzini et al. (2012), use proprietary trade datasets.

6.4.2.1 The Effect of Proportional Transaction Costs

Schultz (1983) and Stoll and Whaley (1983) are among the earliest studies in this line of research and investigate the effect of commissions and spreads on size-based trading strategies. Both papers note that small firms have higher transaction costs due to lower stock prices and higher bid-ask spreads. The estimates for the average round-trip transaction costs for the smallest decile portfolio in Schultz (1983) are about 11.4 % for the NYSE/AMEX stocks in the sample period from 1963 to 1979. The risk-adjusted returns for the small firms are approximately 31 % per year net of transaction costs, suggesting that the transaction costs cannot explain away the higher average returns from small stocks. Stoll and Whaley (1983) consider both the bid-ask spread and commissions and document that the round-trip transaction costs are about 6.8 % for the smallest decile portfolio and 2.7 % for the largest decile portfolio. If round-trip transactions occur only once a year, the size effect remains significant.

A number of later studies investigate the effects of explicit trading costs on prior-return based (momentum and contrarian) trading strategies. Ball, Kothari, and Shanken (1995) show that microstructure effects, such as bid/ask spreads, significantly reduce the profitability of a contrarian strategy. Grundy and Martin (2001) calculate that with round-trip transactions costs of 1.5 %, the profits on a long/short momentum strategy become statistically insignificant. With round-trip transactions costs of 1.77 %, they find that the profits on the long/short momentum strategy are driven to zero.

Novy-Marx and Velikov (2014) study the effect of spread on a broad set of 23 anomalies which vary in terms of the rate of portfolio turnover and are categorized into three distinct bins based on the turnover rate of the portfolios. For example, value strategies are in the low turnover category because the portfolios are rebalanced annually. The strategies in the mid-turnover group have higher rebalancing frequencies including the momentum anomaly, the idiosyncratic volatility anomaly, and the post-earnings-announcement drift. The high-turnover group contains the strategies that involve frequent rebalancing such as the short-term reversal (e.g., Lehmann (1990)) and the seasonality anomaly in Heston and Sadka (2008). Their measure of transaction cost is the effective bid-ask spread measure in Hasbrouck (2009). They find that the low-turnover strategies, typically those using annual rebalancing, survive after taking into account the effect of spread. The transaction costs in such strategies are often less than 10 bps per month. The middle-turnover strategies experience average trading costs between 20 bps and 57 bps per month, typically more than half of these strategies' gross profits. But quite a few of these mid-turnover strategies remain profitable even after considering effective spreads. Trading on the high-turnover anomalies proves to be quite costly and transaction costs wipe out all of the abnormal profits.

6.4.2.2 The Effect of Nonproportional Transaction Cost

Knez and Ready (1996) study the effects of price impact on the profitability of a trading strategy based on the weekly cross-autocorrelations between the returns on large-firm and small-firm portfolios. They find that the trading costs swamp the abnormal returns of the strategy. Mitchell and Pulvino (2001) incorporate commissions and price-impact costs into a merger arbitrage portfolio strategy and find that such costs reduce the profits of the strategy by 300 basis points per year.

Chen et al. (2005)

Chen et al. (2005) implement long-short arbitrage strategies based on the size, book-to-market, or momentum anomaly in the period from 1963 to 2000. They estimate a non-linear concave price-impact function on 5,173 stocks traded on the NYSE and NASDAQ from January 1993 to June 1993. They then estimate the maximal fund size possible before excess returns become negative after taking into account the price impact costs, and find that the profitable fund sizes are small. They therefore conclude that price-impact costs deter agents from exploiting the anomalies and markets are bounded-rational.

Korajczyk and Sadka (2004)

Korajczyk and Sadka (2004) focus on the profitability of long positions in winner-based momentum strategies and incorporate several models of trading costs, including proportional and nonproportional costs. Two proportional cost models are based on quoted and effective spreads. They study two alternative price impact models (nonproportional costs): One based on Glosten and Harris (1988), and one based on Breen et al. (2002).

Since the TAQ data only begin in 1993 and the momentum strategy can be implemented in a much larger time period (from 1967 to 1999), Korajczyk and Sadka (2004) obtain the out-of-sample estimates of transaction costs by (1) estimating cross-sectional regressions of measures of transaction costs obtained from TAQ data on firm-specific variables and (2) applying the time series averages of the estimated coefficients onto the firm-specific variables observable in the out-of-sample period.

Since non-proportional transaction costs, as a percentage of trade size, grow with the size of the portfolio being traded. Their main interest is to determine the size of a single portfolio that will break even with a benchmark. Their main benchmark returns are the Fama-French three-factors. Specifically, they calculate the size of the portfolio that (1) eliminates the statistical significance of the portfolio abnormal return, (2) drives the level of abnormal return to zero, and (3) drives the portfolio Sharpe ratio (Sharpe 1966) to that of the maximal Sharpe ratio obtained from combinations of the Fama and French (1993) market, size, and book-to-market portfolios.

The main sample in Korajczyk and Sadka (2004) contains NYSE stocks from the time period 1967–1999. They find that proportional transaction costs do not have a significant impact on the profitability of momentum strategies. They find that equal-weighted strategies, which are common in the literature, have the highest returns before price impact but the lowest returns after price impact. Price impact quickly eliminates the profitability of equal-weighted portfolios. Value-weighted strategies provide higher post price impact returns than equal-weighted strategies. The break-even fund size for value-weighted strategies is $2 billion—expressed relative to market capitalization at the end of December 1999—using the Breen et al. (2002) price-impact specification. The break-even size is slightly larger with the Glosten and Harris (1988) specification.

In sum, Korajczyk and Sadka (2004) find that commonly cited strategies (equal-weighted or value-weighted) cannot be profitably implemented for large portfolios. They advocate to adopt liquidity-tilted portfolios, which they show perform substantially better than either equal- or value-weighted portfolio, after price impact is taken into account. They conclude, therefore, transaction costs—in the form of spreads and price impacts of trades—do not fully explain the return persistence of past winner stocks exhibited in the data. They deem their break-even funds sizes conservative, as their analysis assumes that trading per stock is completed in a single trade over a 30-min time interval. More sophisticated trading mechanisms that break-up large trades over the trading day may reduce trading costs and may further increase the computed break-even fund sizes.

Frazzini et al. (2012)

Frazzini et al. (2012) examine the net of trading costs returns of the size, value, momentum, and short-term reversal strategies (SMB, HML, UMD, and STR) by using a live trading dataset from a large institutional money manager over the time period from 1998 to 2011 across 19 developed equity markets. The data contain information about orders, execution prices, and quantities traded. The data are quite unique insofar as they contain both the actual trade and the intended trade of the fund manager. The dataset includes $721 billion worth of trades executed using automated algorithms (trades associated with intra-day high frequency models are excluded).

Overall, the transaction cost estimates in Frazzini et al. (2012) are less than a tenth as large as previous studies suggest. At the implied fund size of 15.2 billion dollars, the transaction cost of the SMB strategy is 1.46 % per year. Using the historical average gross return of 2.61 % and the implied price impact function they estimated using actual trading data, they obtain the break-even fund size for the SMB strategy to be $103 billion. The break-even fund sizes are $82.95 billion for HML, $52.15 billion for UMD, and $9.51 billion for STR.

A close look at the cost estimates reveals that they are quite similar in magnitude to those used in Korajczyk and Sadka (2004) over the years for which the sample periods of the two studies overlap, that is, pre-decimalization. Indeed, after decimalization trading costs have substantially declined (this can also be seen in Fig. 6.1 discussed below), suggesting that the change in trading environment makes some trading strategies feasible although they might not have been in the past.

6.4.3 Optimized Portfolios

Most of the strategies studied in the literature are standard long or long-short strategies, without taking into account the actions that fund managers can take to reduce the effect of transaction costs. Strategies that are cost-conscious can substantially enhance performance. Investors can reduce transaction costs by shifting towards stocks with lower transaction costs as well as reducing the turnover rate of portfolios. To improve the net-of-transaction costs performance, one needs to reduce the transaction cost, while still maintaining adequate exposure to the underlying signal.

One can start by limiting trades to liquid stocks only. Novy-Marx and Velikov (2014) find that this approach does not yield significant improvement for mid-turnover strategies, although the improvement is significant for high-turnover strategies. They also propose a staggered partial rebalancing technique which involves lowering the frequency at which a strategy is traded. For example, for the momentum strategy, traders can use quarterly rebalancing instead of monthly. But the effect of the staggered partial rebalancing technique turns out to be quite limited for mid-turnover strategies, probably due to the loss of the exposure to the

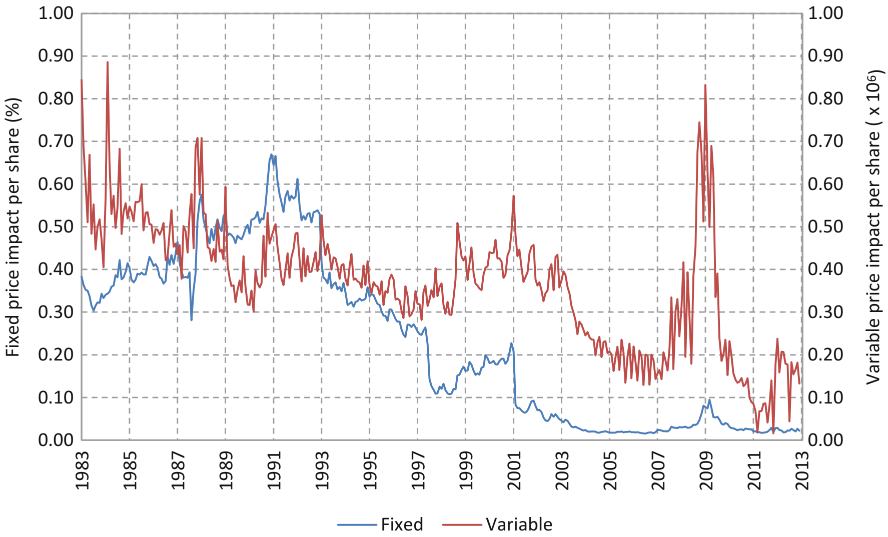


Fig. 6.1 The time series of market liquidity. The figure plots the monthly average of price-impact components estimated as in Sadka (2006). The fixed component measures the transitory price impact per trade, from one trade to the next; the variable component measures the permanent price impact (from one trade to the next) per share traded. Price-impact components per firm are scaled by beginning-of-month share price, thus interpreted as relative cost values. The *vertical grid lines* represent the January month of the year labeled. The universe consists of NYSE-listed firms for the period January 1983 through December 2012

underlying signal. Novy-Marx and Velikov find that a buy-hold strategy that reduces the turnover rate while maintaining the exposure to the underlying signal is more effective. An example in the case of the momentum strategy would be a 10 %/20 % buy/hold rule, which is modified from the traditional decile portfolio approach and involves buying a stock when the stock gets into the top 10 % and holding it unless the stock leaves the top 20 %. The average reduction in the turnover is 41 % with a 42 % decrease in transaction costs. The buy/hold strategy can also improve the performance of the high-turnover strategy.

Korajczyk and Sadka (2004) derive a simple liquidity-weighted portfolio rule that takes into account the price impact costs of trading in exploiting the momentum anomaly. Specifically, they study a partial, static optimization problem and aim to maximize, under simplifying assumptions, post-price impact expected return on the portfolio. They find that liquidity-tilted trading strategies can be implemented for much larger portfolios. Alpha is driven to zero after $5 billion is invested for the liquidity-weighted strategy, or the 50/50 weighting of the liquidity-weighted or value-weighted strategies. Korajczyk and Sadka point out that an extension of the static optimization to a dynamic setting should result in portfolios of even superior performance. Performance might also be improved if some simplifying assumptions are relaxed.

Frazzini et al. (2012) also study how to maximize after-trading cost returns while maintaining the style of the original long-short portfolio. Their baseline optimization problem is to minimize the expected trading cost and in the meantime require that the ex-ante tracking error be less than 1 % and the trading volume be less than 5 % of a stock's average daily trading volume. They find a significant reduction in transaction costs in the anomalies that they study. For example, the trading cost for the standard HML strategy is 2.28 % per year; however, their optimized HML strategy only generates 57 bps of transaction costs per year. Since the optimization is achieved with the restriction on the tracking error, the optimized portfolios, therefore, maintain the styles of the original portfolios.

6.5 Liquidity Over Time

One of the key developments in the US equity market over the past several decades is the substantial increase in institutional investing and index trading. It is well accepted that increases in institutional investing and index trading have played a key role in the increases of trading volume and liquidity levels in US equity markets. The discussion of profitability under transactions costs therefore cannot be complete without highlighting the time series dimension.

To highlight the reduction in trading costs, Fig. 6.1 plots the time series of the average price impact measures calculated in Sadka (2006) for the period 1983–2012. Sadka follows the estimation procedure in Glosten and Harris (1988), with adjustments for the potential autocorrelation in the order flow as well as the differential impact of large trades. The fixed component of price impact displays a significant drop in magnitude of the sample period; it exhibits significant declines around the reductions of tick size (1/8 to 1/16 in June 1997 and to decimals in January 2001). The level of the fixed component has been stable at a low level since 2001, except for a slight increase during the financial crisis of 2008–2009. The variable component of price impact has also declined over time, but at a slower rate. The financial-crisis period shows high levels of price impact, similar to its levels during the early 1980s, but they drop back to pre-crisis levels in 2010.

The drop in transaction costs would suggest an increase in profitability of trading strategies—that is, strategies that would have seem unprofitable under trading costs in the past, might seem more attractive under the current, low cost regime. However, there are two noteworthy consequences of the increased liquidity over the years: (a) increased liquidity commonality and (b) decreased mispricing.

Liquidity Commonality Chordia et al. (2000) document that the time series of liquidity across a large panel of stocks displays a significant systematic component, that is, there exists a significant correlation between the liquidity of different firms over time (see also Hasbrouck and Seppi (2001) and Korajczyk and Sadka (2008)). This suggests the existence of liquidity risk, which is the risk of experiencing a sudden drop in market liquidity—a risk not easily diversifiable across the different stocks in a portfolio of a given investment manager. Kamara et al. (2008) show that liquidity commonality has substantially increased over time, especially for large-cap firms. They show a causal relation between the increase in institutional investing and liquidity commonality over the period 1963–2005. Lou and Sadka (2011) further highlight the difference between the liquidity level (as measured by the Amihud (2002) measure) and the liquidity risk (as measured in Pástor and Stambaugh (2003) and Sadka (2006)) of a given firm, and show that during the crisis period of 2008–2009 the liquidity risk subsumed liquidity level in explaining return differences in the cross-section of stocks. In fact, some large, liquid firms have exhibited high liquidity risk. Therefore, the conclusion is that although transactions costs have decreased over time, thereby making the post-transaction-costs expected returns more attractive, the trend might also be associated with an increase in the liquidity risk of the investment strategies designed to capture asset-pricing anomalies. This is of special concern for investment strategies that rely on a substantial amount of trading because their performance is more exposed to systematic changes in market liquidity.

Mispricing The informativeness of prices is directly affected by the costs of trading (for recent works see Sadka and Scherbina (2007) and Kerr et al. (2014)). Therefore, one might expect the ex-ante profitability of anomaly-based investment strategies to decline, as more traders are prone to trade them given the reduction in trading costs. An analysis of the profitability of prominent trading strategies, such as size, value, price and earnings momentum, and idiosyncratic volatility, over time is provided in Chordia et al. (2014). They show that, without considering transactions costs, the "paper" profits of such popular strategies have significantly declined over time. They attribute this decline in profitability to the increased arbitrage activity and improvements in market liquidity, and offer this evidence in support of capital market efficiency. Evidence for increased trading activity is also seen by the decrease in the average investment horizon of institutional investors over the past several decades. Figure 6.2 plots the time series of the average Churn ratio; this ratio is based on the turnover in institutional investors quarterly holdings gauged from the 13F filings (see Gaspar et al. (2005) and Kamara et al. (2016)). The plot suggests that institutional investors have increased portfolio turnover, possibly focusing on shorter horizon investment strategies, which can be more attractive in the current trading environment. Yet, this increased sensitivity of profitability to the costs of trading coupled with the aforementioned increase in liquidity risk makes these investment portfolios riskier.

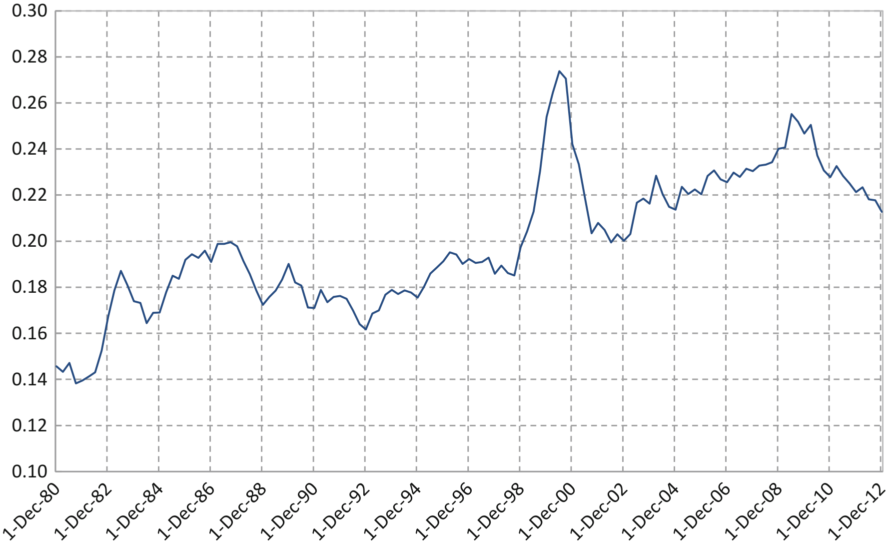


Fig. 6.2 The time series of portfolio turnover ratio of institutional investors. The figure plots the quarterly institution turnover ratio estimated as in Kamara, Korajczyk, Lou, and Sadka (2016). For each quarter q and institution j from 1980 to 2012, the churn ratio is defined as:

$$Churn_{j,q} = \frac{\Sigma|Shares_{i,j,q}P_{i,q} - Shares_{i,j,q-1}P_{i,q-1} - Shares_{i,j,q-1}\Delta P_{i,q}|}{\Sigma_i(Shares_{i,j,q}P_{i,q} + Shares_{i,j,q-1}P_{i,q-q})/2},$$

where $Shares_{i,j,q}$ is the number of shares of firm i owned by institution j at the end of quarter q and $P_{i,q}$ is the price of stock i at the end of quarter q. The institution portfolio turnover ratio is calculated as the average churn ratio in the four quarters [q−3,q] for each j and q. Plotted is the time series of the average turnover ratio across all institutions in the 13-F dataset from 1980 to 2012

6.6 Conclusion

This article discusses the various implications of transactions costs for the profitability of investment strategies. In his many papers in this area, Treynor has substantially contributed to our understanding of the sources of trading costs and their implications for portfolio management. With the recent advances in trading systems and the need to use complex trading algorithms to reduce institutional investors fear of "being skimmed" by high-frequency traders, the concept of the invisible costs of trading remains a significant and relevant topic of discussion amongst academics and practitioners.

Acknowledgements We would like to thank Robert Korajczyk for helpful comments and Siyi Shen for valuable research assistance.

References

Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: volume and price variability. *The Review of Financial Studies, 1*, 3–40.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets, 5*, 31–56.

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance, 68*(3), 929–985.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*, 159–178.

Ball, R., Kothari, S. P., & Shanken, J. (1995). Problems in measuring portfolio performance: An application to contrarian investment strategies. *Journal of Financial Economics, 38*, 79–107.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*, 3–18.

Breen, W. J., Hodrick, L. S., & Korajczyk, R. A. (2002). Predicting equity liquidity. *Management Science, 48*(4), 470–483.

Bhushan, R. (1994). An informational efficiency perspective on the post-earnings announcement drift. *Journal of Accounting and Economics, 18*(1), 45–65.

Chakrabarty, B., Li, B., Nguyen, V., & Van Ness, R. A. (2007). Trade classification algorithms for electronic communications network trades. *Journal of Banking and Finance, 31*(12), 3806–3821.

Chakrabarty, B., Bingguang, L., Van Thuan, N., & Van Ness, R. (2007). Trade classification algorithms for electronic communications networks. *Journal of Banking and Finance, 31*, 3806–3821.

Chan, L. K. C., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *Journal of Finance, 51*, 1681–1713.

Chan, L. K. C., Jegadeesh, N., & Lakonishok, J. (1999). The profitability of momentum strategies. *Financial Analysts Journal, 55*, 80–90.

Chen, Z., Stanzl, W., & Watanabe, M. (2005). Price impact costs and the limit of arbitrage. Working paper, Yale School of Management.

Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics, 56*, 3–28.

Chordia, T., Goyal, A., Sadka, G., Sadka, R., & Shivakumar, L. (2009). Liquidity and the post-earnings-announcement drift. *Financial Analysts Journal, 65*, 18–32.

Chordia, T., Subrahmanyam, A., & Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Aclcounting and Economics, 58*, 41–58.

Corwin, S. A., & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance, 67*, 719–760
Chakrabarty, Bidisha, Bingguang Li, Van Thuan Nguyen, and Robert Van Ness, 2007, Trade classification algorithms for Electronic Communications Networks, Journal of Banking and Finance 31, 3806-3821.

Easley, D., & O'hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics, 19*, 69–90.

Ellis, K., Roni, M., & O'hara, M. (2000). The accuracy of trade classification rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis, 35*, 529–551.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383–417.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Financial Economics, 47.2*, 427–465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*, 55–84.

Fama, E. F., & French, K. R. (2010). Luck versus Skill in the cross-section of mutual fund returns. *Journal of Finance, 65*, 1915–1947.

Fialkowski, D., & Petersen, M. A. (1994). Posted versus effective spreads: Good prices or bad quotes? *Journal of Financial Economics, 35*, 269–292.

Frazzini, A., Israel, R., & Moskowitz, T. J. (2012). Trading costs of asset pricing anomalies. Unpublished working paper. University of Chicago.

Gaspar, J. -M., Massab, M., & Matos, P. (2005). Shareholder investment horizons and the market for corporate control. *Journal of Financial Economics, 76*, 135–165.

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*, 71–100.

Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics, 21.1*, 123–142.

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*, 153–181.

Graham, B., & Dodd, D. L. (1934). *Security analysis*. McGraw-Hill Book Co. https://en.wikipedia.org/wiki/Security_Analysis_(book).

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review, 70*(3), 393–408.

Grundy, B. D., & Martin, J. S. (2001). Understanding the nature of the risks and the sources of the rewards to momentum investing. *Review of Financial Studies, 14*, 29–78.

Hasbrouck, J. (1991a). The summary informativeness of stock trades: an econometric analysis. *Review of Financial Studies, 4*(3), 571–595.

Hasbrouck, J. (1991b). Measuring the information content of stock trades. *Journal of Finance, 26*, 179–207.

Hasbrouck, J., & Seppi, D. (2001). Common factors in prices, order flows and liquidity. *Journal of Financial Economics, 59*, 383–411.

Hasbrouck, J. (2009). Trading costs and returns for US equities: estimating effective costs from daily data. *Journal of Finance, 64*, 1445–1477.

Huberman, G., & Stanzl, W. (2004). Price Manipulation and quasi-arbitrage. *Econometrica, 72*, 1247–1275.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48.1*, 65–91.

Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance, 56.2*, 699–720.

Kamara, A., Lou, X., & Sadka, R. (2008). The divergence of liquidity commonality in the cross-section of stocks. *Journal of Financial Economics, 89*, 444–466.

Kamara, A., Korajczyk, R. A., Lou, X., & Sadka, R. (2016). Horizon pricing. Unpublished working paper.

Kamara, A., Korajczyk, R. A., Lou, X., & Sadka, R. (2016). Horizon pricing. Journal of Financial and Quantitative Analysis, forthcoming.

Keim, D. B., & Madhavan, A. (1996). The upstairs market for large-block transactions: analysis and measurement of price effects. *Review of Financial Studies, 9*, 1–36.

Keim, D. B., & Madhavan, A. (1997). Transactions costs and investment style: An inter-exchange analysis of institutional equity trades. *Journal of Financial Economics, 46*, 265–292.

Kerr, J. N., Sadka, G. & Sadka, R. (2014). Illiquidity and earnings predictability. Boston College, working paper.

Knez, P. J. & Ready, M. J. (1996). Estimating the profits from trading strategies. *Review of Financial Studies, 9*, 1121–1163.

Korajczyk, R. A. & Sadka, R. (2004). Are momentum profits robust to trading costs? *The Journal of Finance, 59*, 1039–1082.

Korajczyk, R. A., & Sadka, R. (2008). Pricing the commonality across alternative measures of liquidity. *Journal of Financial Economics, 87*, 45–72.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 1315–1335.
Lee, C. M. C. & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance, 46*, 733–754.
Lehmann, B. (1990). Fads, martingales and market efficiency. *Quarterly Journal of Economics, 105*, 1–28.
Lesmond, D. A., Schill, M. J. & Zhou, C. (2004). The illusory nature of momentum profits. *Journal of Financial Economics, 71.2*, 349–380.
Loeb, T. F. (1983). Trading cost: The critical link between investment information and results. *Financial Analysts Journal, 1983*, 39–44.
Lou, X. & Sadka, R. (2011). Liquidity level or liquidity risk? Evidence from the financial crisis. *Financial Analysts Journal, 67*, 51–62.
McLean, D. R. & Pontiff, J. (2014). Does Academic Research Destroy Stock Return Predictability? Boston College working paper.
Mendenhall, R. R. (2004). Arbitrage risk and post-earnings-announcement drift. *The Journal of Business, 77*, 875–894.
Mitchell, M. & Pulvino, T. (2001). Characteristics of risk and return in risk arbitrage. *Journal of Finance, 56*, 2135–2175.
Ng, J., Rusticus, T. O., & Verdi, R. S. (2008). Implications of transaction costs for the post-earnings-announcement drift. *Journal of Accounting Research, 46*, 661–696.
Novy-Marx, R. & Velikov, M. (2014). A Taxonomy of Anomalies and their Trading Costs, NBER working paper No. 20721.
Pástor, Ľ. & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*, 642–685.
Sadka, R. (2006). Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics, 80*, 309–349.
Sadka, R. & Scherbina, A. (2007). Analyst disagreement, mispricing, and liquidity. *Journal of Finance, 62*, 2367–2403.
Schultz, P. (1983). Transactions costs and the small firm effect: A comment. *Journal of Financial Economics, 12*, 81–88.
Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 39*, 119–138.
Stattman, D. (1980). Book values and stock returns. *Chicago MBA: A Journal of Selected Papers, 4*, 25–45.
Stoll, H. R. & Whaley, R. E. (1983). Transactions costs and the small firm effect. *Journal of Financial Economics, 12*, 57–80.
Treynor, J. (1971). The only game in town. *Financial Analysts Journal, 27*, 12–14.
Treynor, J. (1981). What does it take to win the trading game? *Financial Analysts Journal*, 55–60.
Treynor, J. (1987). The economics of the dealer function. *Financial Analysts Journal*, 27–34.
Treynor, J. (1994). The invisible cost of trading. *The Journal of Portfolio Management, 21*, 71–78.

Chapter 7
Mean-ETL Portfolio Construction in US Equity Market

Barret Pengyuan Shao

7.1 Introduction

Fundamental data have been used as the criteria to select stock for a long time in both industry and academic work. For example, Jaffe, Keim, and Westerfield (1989) find the relation between stock returns and the earnings to price ratio to be significant. Fama and French (1992) add fundamental data of stocks, size and book-to-market to explain stock returns. Jegadeesh and Titman (1993) show the price momentum effect that buying stocks that performed well in the past and selling stocks that performed poorly in the past can generate statistically significant positive returns over both 3-month and 12-month holding periods. Bloch, Guerard, Markowitz, Todd, and Xu (1993) develop an underlying composite model to describe stock returns using fundamental variables. By adding the consensus earnings forecasting (CTEF) and price momentum (PM) variables to this model, Guerard, Gültekin, and Stone (1997) develop the US Expected Returns (USER) stock selection model.

As the pioneering work, Markowitz (1952 1959) uses a mean-variance portfolio construction model to maximize the portfolio return for a given level of risk. In Markowitz's mean-variance portfolio optimization, the portfolio risk is represented by the portfolio variance. There are many previous works that have been done to examine the efficiency of mean-variance portfolio on fundamental variables in the stock markets. For example, Guerard, Rachev, and Shao (2013) show that both mean-variance and mean-ETL portfolios on global expected returns (GLER) model are capable of generating statistically significant active returns.

The organization of this chapter is as follows, Sect. 7.2 provides an overview of the importance of different fundamental variables used in the mean-ETL portfolio construction. Section 7.3 describes the methodology used to generate the scenarios

B.P. Shao (✉)
Crabel Capital Management LLC, Charlottesville, VA 22903, USA
e-mail: barretshao@gmail.com

© Springer International Publishing Switzerland 2017
J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_7

and construct the portfolios based on the fundamental variables. Simulation results of the mean-ETL portfolios during the period 1990–2013 are presented in Sect. 7.4 and we provide the conclusions in Sect. 7.5.

7.2 Fundamental Variables

In this section, we discuss different fundamental variables that are used to construct the US mean-ETL portfolio in this study which are consensus temporary earnings forecasting (CTEF), price momentum (PM), US expected return (USER) and McKinley Capital Quant Score (MQ) variables.

As one of the most important fundamental variables, earning expectation has been used for stock selection for many years. Graham, Dodd, and Cottle (1934) select stocks based on fundamental valuation techniques, and show that stocks with higher earnings-per-share (EPS) outperform the ones with lower EPS. In 1975, Lynch, Jones, and Ryan collected and published consensus statistics for EPS forecasting, forming the beginning of what is now known as the Institutional Brokerage Estimation Service (I/B/E/S) database (Brown 1999). Besides the original earnings yield, Guerard et al. (1997) also find that the earnings revision (EREV) and earnings breadth (EB) which represents the direction of the revisions are also important in stock selection. Guerard et al. (1997) create CTEF variable which is a weighted sum of the forecasted earnings yield (FEP) from I/B/E/S, EREV, and EB. Guerard, Blin, and Bender (1998) find that a value-based model with CTEF variables produces statistically significant models for stock selection in the US and Japanese markets. More recently, using the CTEF variables of global equities, Xia, Min, and Deng (2015) show that the mean-variance portfolio construction produces robust returns. Similarly, also using the CTEF variables of global equities, Shao, Rachev, and Mu (2015) show that robust returns can be generated via the mean-ETL portfolio construction.

In the global expected return model (see Guerard et al. 2013), CTEF and PM variables account for the majority of the forecast performance. The role of PM variable in stock selection has also been studied for a long time (Jegadeesh & Titman 1993) find that 3-month, quarterly, 6-month, and 1-year PM variables are statistically significant with excess returns. Grinblatt, Titman, and Wermers (1995) find that 77 % of the mutual funds are buying stocks with better performance history. Chan, Hameed, and Tong (2000) report that trading strategies based on PM in international equity markets generate statistically significant returns. Brush (2001) shows that quarterly information coefficient (IC) of 3-month PM variable is 0.073, which is higher than the monthly IC of 0.053. The PM variable in this study is defined as the price from last month divided by the price from12 months prior (see Guerard et al. 2013, for details). However, implementing portfolio construction using PM variable on the US or global markets has not been studied before. We examine the mean-ETL-PM portfolio construction on US market in this chapter. We also test the same portfolio construction technique on the McKinley Capital

Quant Score (MQ) variable which is the equally weighted sum of CTEF and PM variables.

By adding CTEF and PM variables to the original stock selection model developed by Bloch et al. (1993), Guerard et al. (1997) construct the USER model based on fundamental data in US market. The USER variable we use in this study to construct the portfolio is the same as the one used in the USER model developed by Guerard et al. (1997). At time $t+1$, the USER variable is a weighted sum of certain fundamental variables and their derivatives at time t. These fundamental variables and their derivatives include:

1. earnings—price ratio (EP);
2. book value—price ratio (BP);
3. cash flow—price ratio (CP);
4. sales—price ratio (SP);
5. current EP ratio divided by average EP ratio over the last 5 years (REP);
6. current BP ratio divided by average BP ratio over the last 5 years (RBP);
7. current CP ratio divided by average CP ratio over the last 5 years (RCP);
8. current SP ratio divided by average SP ratio over the last 5 years (RSP);
9. csensus temporary earnings forecasting (CTEF);
10. price momentum (PM).

For more details on the methodology and estimation of the USER model, the reader is referred to Guerard et al. (2013). Guerard et al. (2013) also report the mean-variance portfolio using USER data can generate an active annual return of 10.70 % with an information ratio of 1.12 and a t-statistic of 3.68 during the period 1999–2009. In this chapter, the mean-ETL portfolio on the USER variable during the period 2000–2013 is examined.

7.3 Mean-ETL Portfolio Construction

In this section, we describe the mean-ETL portfolio construction methodology used in this chapter. Markowitz's mean-variance portfolio optimization (1952, 1959) is to maximize the portfolio expected return at a given level of portfolio risk. In Markowitz's framework, the portfolio expected return is measured by the sum of the security weights multiplied by their respective expected return, and portfolio risk is measured by the variance of portfolio. Instead of using the portfolio variance as the risk measure, mean-ETL portfolio optimization uses *expected-tailed loss* (ETL), also known as *conditional Value-at-Risk* (CVaR) or the *expected shortfall* (ES), as the risk measure. Many recent research works show the mean-ETL portfolio construction can generate robust portfolios. For example, Shao and Rachev (2013) show that the mean-ETL portfolio construction based on GLER variable generate statistically significant active return on the global markets.

7.3.1 Mean-ETL Framework

Before introducing the definition of ETL, we need to give the definition of value-at-risk (VaR), which is one of the most frequently used risk measures in the finance industry:

Let X represent the distribution of portfolio returns; then, the VaR of the portfolio at a $(1-\alpha)\,100\,\%$ confidence level can be defined as the lower α quantile of the return distribution:

$$\mathrm{VaR}_{\alpha}(\mathrm{X}) = -\inf\left(x : \mathrm{P}\left(\mathrm{X} \le x\right) \ge \alpha\right) = -\mathrm{F}_{X}^{-1}(\alpha) \tag{7.1}$$

Given the definition of VaR, ETL is defined as the average loss beyond the VaR threshold:

$$\mathrm{ETL}_{\alpha}(X) = -E\left(X \middle| X < -\mathrm{VaR}_{\alpha}(X)\right) \tag{7.2}$$

As a coherent risk measure (Artzner, Delbaen, Eber, & Heath 1999), ETL has many good properties for the purpose of portfolio risk management. Rockafellar and Uryasev (2000 2002) provide a detailed discussion of the properties of CVaR when used as a measure of risk. The main reasons that we are using ETL instead of VaR as the risk measure in portfolio construction are summarized by Rachev, Martin, Racheva, and Stoyanov (2009): (1) ETL gives a more informative view on extreme events; (2) mean-ETL portfolio optimization problem is a convex problem, which has a unique solution; (3) as a coherent risk measure with the sub-additivity property, ETL accounts for the effect of diversification. For more details on the calculation of ETL in portfolio optimization, the reader is referred to Shao et al. (2015).

Using ETL as the risk measure, there are several ways to formulate the mean-ETL optimization problem (see Rachev, Stoyanov, and Fabozzi (2007) for details). In this study, we add a few constraints to the portfolio construction: (1) the portfolio needs to be a long only portfolio; (2) maximum weights of single asset cannot exceed 4 %; (3) monthly turnover rate of the portfolio is below or equal to 8 %. We formalize our mean-ETL optimization framework as follows:

$$\begin{aligned} &\max_{\mathbf{w_t}} \tfrac{1}{S}\sum_{s=1}^{S}\mathbf{w}_t^{\mathrm{T}}\hat{\mathbf{Y}}_t^{(s)} - \lambda \mathrm{ETL}_{\alpha}\left(\mathbf{w}_t^{\mathrm{T}}\hat{\mathbf{Y}}_t\right) \\ &\text{s.t.}\quad 0.04 \ge \mathbf{w}_t \ge 0 \\ &\mathbf{e}_1^{\mathrm{T}}\left|\mathbf{w}_t - \mathbf{w}_{t-1}\right| \le 0.16 \\ &\mathbf{e}_1^{\mathrm{T}}\mathbf{w}_t = 1 \end{aligned} \tag{7.3}$$

where $\mathbf{w}_t$ is a column vector of optimal securities weights in the portfolio and $\hat{\mathbf{Y}}_t^{(s)}$ is a column vector that contains the sth scenario of all N securities:

$$\hat{\mathbf{Y}}_t^{(s)} = \left[\hat{Y}_t^{(1,s)}, \hat{Y}_t^{(2,s)}, \ldots, \hat{Y}_t^{(N,s)}\right]^{\mathrm{T}},$$

and $\hat{\mathbf{Y}}_t$ should be a matrix that consists of all S scenarios at time t:

$$\hat{\mathbf{Y}}_t = \left[\hat{\mathbf{Y}}_t^{(1)}, \hat{\mathbf{Y}}_t^{(2)}, \ldots, \hat{\mathbf{Y}}_t^{(S)}\right]_{N \times S}.$$

In this study of mean-ETL portfolio construction, α for ETL is chosen to be 5 % and the risk-averse parameter λ is chosen to be 1.

7.3.2 Scenario Generator

With the above-mentioned mean-ETL portfolio construction framework, the scenarios of fundamental variables, $\hat{\mathbf{Y}}_t$, are required for the optimization.

In this study, we use $Y_{n,t}$ to denote fundamental variable values of the n^{th} stock at time t, where $1 \le t \le T,\ \ 1 \le n \le N$. Here, T and N are the number of data points and the number of securities, respectively. The fundamental variable is chosen from USER, CTEF, PM, and MQ in this study. Before fitting the time series model on the fundamental variable data, we perform the following transformations for the n^{th} stock:

$$y_{n,t} = f(x) = \begin{cases} \log\left(Y_{n,t}\right) - \log\left(Y_{n,t-1}\right), \; t \ge 2 \quad \text{and} \quad k = 1, \ldots, N \\ 0, \qquad\qquad\qquad\qquad\qquad t = 1 \quad \text{and} \quad k = 1, \ldots, N, \end{cases} \tag{7.4}$$

where N denotes the number of securities in the portfolio.

We choose the time series model autoregressive moving average (ARMA) generalized autoregressive conditional heteroskedastic (GARCH) with multivariate normal tempered stable (MNTS) innovations to model and generate the scenarios of time series $y_{n,t}$. The ARMA-GARCH-MNTS model is a flexible model to capture volatility clustering, heavy tails, asymmetric dependence structure and the dependency between different stocks. The AMRA-GARCH-MNTS model used to model time series $y_{n,t}$ in this study is described as follows:

$$y_{n,t} = c_n + a_n y_{n,t-1} + b_n \sigma_{n,t} \eta_{n,t-1} + \sigma_{n,t} \eta_{n,t} \tag{7.5}$$

$$\sigma_{n,t}^2 = \alpha_{n,0} + \alpha_{n,1} \sigma_{n,\,t-1}^2 \eta_{n,t-1}^2 + \beta_{n,1} \sigma_{n,t-1}^2, \tag{7.6}$$

where $n = 1, 2, \cdots, N$ and the joint innovation term $\boldsymbol{\eta_t} = (\eta_{1,t}, \eta_{2,t}, \cdots, \eta_{N,t})$ is generated from $MNTS(\alpha, \theta, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\rho})$. And the N-dim $MNTS(\alpha, \theta, \beta, \gamma, \mu, \rho)$ is defined as:

$$X = (X_1, X_2, \ldots, X_N) = \boldsymbol{\mu} + \boldsymbol{\beta}\,(CT - 1) + \boldsymbol{\gamma}\sqrt{CT}\boldsymbol{\varepsilon} \tag{7.7}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)^{\mathrm{T}}$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_N)^{\mathrm{T}}$, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_N)^{\mathrm{T}}$ and $\boldsymbol{\gamma} > 0$. CT is a classical tempered stable (CTS) subordinator with parameter (α, θ), where $\alpha \in (0, 1)$ and $\theta > 0$. $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N)^{\mathrm{T}}$ is a N-dim standard normal distribution with covariance matrix $\boldsymbol{\rho}$, independent of the subordinator $CT > 0$. The reader is referred to Kim, Giacometti, Rachev, Fabozzi and Mignacca (2012) for the details on the application and estimation of MNTS in portfolio optimization.

In this chapter, we are not going into the details surrounding the estimation of ARMA-GARCH-MNTS model and the reader is referred to Shao et al. (2015) for the estimation methodology. With the estimated ARMA-GARCH-MNTS model, we can generate the scenarios of $y_{n,t}$, denoted as $\hat{y}^s_{n,t}$ and it represents the sth scenario forecasting the nth stock at time t. To get the scenarios of the fundamental variables $R_{n,t}$, we do the following reverse transformation:

$$\hat{Y}^s_{n,t} = \exp\left(\hat{y}^s_{n,t} + \log\left(Y_{n,t-1}\right)\right) \tag{7.8}$$

After generating the scenarios $\hat{Y}^s_{n,t}$ for each stock, we can run the portfolio optimization on different fundamental variables from US markets. The portfolio results with empirical data will be discussed in the next section.

7.4 Portfolio Results and Analysis

We generate the optimal mean-ETL portfolios during the period 2000–2013 based on the scenarios generated from ARMA-GARCH-MNTS. In this chapter, we use mean-ETL-CTEF, mean-ETL-USER, mean-ETL-MQ, and mean-ETL-PM to represent the optimal portfolios based on CTEF, USER, MQ, and PM variable, respectively. In this section, we use the Barra risk model to analyze these portfolios. The US Barra risk model was developed in Rosenberg and Marathe (1975) based on company fundamental data. The Barra attribution analyses here use the US Equity Model (USE3). The benchmark in these attribution reports is chosen to be Russell 3000 Growth index. We also offer the comparison of these portfolios in this section.

7.4.1 Attribution Reports

The attribution of mean-ETL-CTEF portfolio returns is shown in Table 7.1, and we report that the portfolio generates 14.07 % active (excess) annual return with an information ratio of 1.11 and a t-statistic of 4.15. The active return is highly statistically significant. The attribution report also shows that the stock selection produces statistically significant active annual return of 5.95 % with a t-statistic of 3.66. Moreover, the earnings yield contributes 1.24 % annual return with a t-statistic of 3.31. The mean-ETL-CTEF portfolio is based on the CTEF variable

Table 7.1 Barra attribution of the mean-ETL-CTEF (US) portfolio

Source of return	Contribution (%)	Avg exposure	Risk (%)	IR	*t*-stat
Portfolio	**16.12**		**20.87**		
Benchmark	**2.05**		**18.30**		
Total active	**14.07**		**12.10**	**1.11**	**4.15**
Expected active	0.09				
Market timing	3.14		5.90	0.47	1.74
Risk Indices	6.07		7.88	0.73	2.73
Volatility	−0.31	0.31	2.67	−0.09	−0.33
Momentum	0.53	−0.19	2.99	0.19	0.70
Size	6.24	−1.64	7.15	0.81	3.01
Size non-linearity	−1.24	−0.68	3.53	−0.33	−1.23
Trading activity	−0.13	−0.21	0.99	−0.12	−0.46
Growth	0.17	−0.42	1.10	0.15	0.56
Earnings yield	1.24	0.10	1.23	0.89	3.31
Value	−0.13	0.68	1.87	−0.03	−0.12
Earnings variation	−0.04	0.30	0.84	−0.03	−0.10
Leverage	−0.30	0.39	1.16	−0.20	−0.73
Currency sensitivity	0.22	−0.26	0.78	0.25	0.92
Yield	−0.04	0.15	0.57	−0.06	−0.22
Non-EST universe	−0.14	0.12	1.27	−0.11	−0.41
Industries	0.54		4.68	0.12	0.46
Asset selection	5.95		5.87	0.98	3.66
Transaction cost	−1.73				

which represents the consensus earnings forecasting and revisions. The statistically significant contribution to the active returns from earnings yield shows the CTEF's predictable ability of earnings.

In Table 7.2, we show the attribution of mean-ETL-USER portfolio returns. The mean-ETL-USER portfolio generates 9.73 % active annual return. The active return is highly statistically significant with a *t*-statistic of 4.26 and it has an information ratio of 1.14 which is higher than the mean-ETL-CTEF portfolio. We find that the size factor has a factor return of 5.49 % with a t-statistic of 4.05. It is also observed that the portfolio's average exposure to the size factor is negative, indicating the mean-ETL-USER portfolio's preference for stocks with small capitalization. Xia et al. (2015) also have the similar observations from the mean-variance portfolio based on USER variable.

The attribution of mean-ETL-MQ portfolio returns is shown in Table 7.3 and it also produces highly statistically significant active return, 10.75 % annually. Moreover, asset selection produces 6.31 % active annual return, which has an information ratio of 1.36 and a t-statistic of 5.07. Both active return and information ratio of asset selection's contribution is the highest among all the four portfolios under comparison.

Table 7.2 Barra attribution of the mean-ETL-USER (US) portfolio

Source of return	Contribution (%)	Avg exposure	Risk (%)	IR	t-stat
Portfolio	**11.78**		**21. 84**		
Benchmark	**2.05**		**18.30**		
Total active	**9.73**		**8.67**	**1.14**	**4.26**
Expected active	0.34				
Market timing	0.33		2.62	0.19	0.70
Risk indices	5.57		5.41	1.00	3.71
Volatility	0.05	0.40	2.07	0.04	0.14
Momentum	0.15	0.01	1.08	0.15	0.55
Size	5.49	−1.52	4.81	1.08	4.05
Size non-linearity	−0.85	−0.75	2.65	−0.30	−1.11
Trading activity	0.16	−0.06	0.25	0.60	2.25
Growth	0.14	−0.24	0.54	0.26	0.97
Earnings yield	1.22	0.16	0.86	1.26	4.70
Value	−0.14	0.57	1.15	−0.08	−0.30
Earnings variation	−0.31	0.36	0.69	−0.36	−1.34
Leverage	−0.145	0.39	0.74	−0.16	−0.58
Currency sensitivity	−0.01	−0.04	0.38	−0.02	−0.09
Yield	0.05	−0.04	0.18	0.28	1.04
Non-EST universe	−0.24	0.20	1.48	−0.16	−0.59
Industries	0.51		3.84	0.18	0.68
Asset selection	5.02		4.50	1.09	3.99
Transaction cost	−2.04				

In Table 7.4 presents the attribution of mean-ETL-PM portfolio returns is presented. The mean-ETL-PM portfolio generates active annual return of 8.51 %. The active return is statistically significant and it has an information ratio of 0.86 and a t-statistic of 3.20.

All of the four portfolios produce statistically significant active returns and they are due mostly to the asset selection. As described previously, α for mean-ETL portfolio optimization in this study is chosen to be 5 % and we want to know whether the technique can control the 5 % left tails well. Taking mean-ETL-CTEF and mean-USER-ETL portfolios as examples, we show the quantile–quantile plots of the mean-ETL portfolio monthly returns against benchmark's in Fig 7.1. The left plot is the mean-ETL-CTEF portfolio against benchmark and the right one is the mean-ETL-USER portfolio against the benchmark. We find that the left tail of mean-ETL portfolio returns is thinner than the benchmark's in most cases. Another interesting observation is that the return distributions of mean-ETL portfolios have even heavier right tails than the benchmark. It indicates that controlling the tail risk of portfolio does not necessarily deteriorate the right tails in this study.

Table 7.3 Barra attribution of the mean-ETL-MQ (US) portfolio

Source of return	Contribution (%)	Avg exposure	Risk (%)	IR	*t*-stat
Portfolio	**12.80**		**22.36**		
Benchmark	**2.05**		**18.30**		
Total active	**10.75**		**9.72**	**1.12**	**4.18**
Expected active	0.83				
Market timing	1.14		4.07	0.33	1.22
Risk indices	4.88		6.22	0.75	2.81
Volatility	−0.39	0.57	3.22	−0.11	−0.43
Momentum	0.10	0.01	1.35	0.10	0.37
Size	5.75	−1.70	5.03	1.08	4.05
Size non-linearity	−0.89	−0.86	2.94	−0.30	−1.10
Trading activity	0.02	−0.06	0.28	0.04	0.16
Growth	0.08	−0.20	0.50	0.18	0.67
Earnings yield	0.96	−0.01	0.99	0.83	3.11
Value	−0.23	0.54	1.16	−0.15	−0.54
Earnings variation	−0.26	0.48	0.98	−0.21	−0.77
Leverage	−0.35	0.42	0.85	−0.33	−1.24
Currency sensitivity	0.12	−0.12	0.33	0.33	1.22
Yield	0.14	−0.12	0.38	0.33	1.25
Non-EST universe	−0.17	0.21	1.67	−0.11	−0.39
Industries	−0.34		3.87	−0.04	−0.14
Asset selection	6.31		4.47	1.36	5.07
Transaction cost	−2.07				

7.4.2 Comparison

Figure 7.2 below shows the equity curve of mean-ETL portfolios and the benchmark portfolio during the period 2000–2013. It's obvious that the mean-ETL-CTEF portfolio performs better than the other portfolios during this period.

To further compare these portfolios in a more quantitative way, we review their Treynor ratios and Sharpe ratios. Treynor (1965) proposes to use beta as the measurement of volatility, which is well known as Treynor ratio, to evaluate the performance of fund. Treynor ratio is calculated as $T = \frac{R - R_f}{\beta}$, where R is the portfolio return, R_f is the risk free rate, and β is the portfolio's beta. One year later, Sharpe (1966) used standard deviation of returns as the measurement of volatility instead of using beta. Table 7.5 presents the performance measured by different risk metrics of the four mean-ETL portfolios under comparison in this study.

All metrics consistently show that the mean-ETL-CTEF portfolio has the best performance among the four portfolios under comparison. In particular, we find that the mean-ETL-CTEF portfolio generates the highest returns while maintains the lowest volatility measured by standard deviation and β. Figure 7.3 presents the

Table 7.4 Barra attribution of the mean-ETL-PM (US) portfolio

Source of return	Contribution (%)	Avg Exposure	Risk (%)	IR	*t*-stat
Portfolio	**10.56**		**25.03**		
Benchmark	**2.05**		**18.30**		
Total active	**8.51**		**11.16**	**0.86**	**3.20**
Expected active	1.16				
Market timing	−0.99		4.19	−0.06	−0.23
Risk Indices	5.15		7.15	0.71	2.65
Volatility	−0.19	0.72	3.68	−0.04	−0.15
Momentum	0.22	0.14	2.00	0.14	0.51
Size	5.66	−1.83	5.43	1.01	3.76
Size non-linearity	−0.73	−1.01	3.48	−0.21	−0.80
Trading activity	0.24	−0.06	0.49	0.46	1.72
Growth	0.04	−0.10	0.40	0.11	0.41
Earnings yield	0.55	−0.06	0.71	0.65	2.41
Value	−0.18	0.50	1.08	−0.12	−0.46
Earnings variation	−0.22	0.54	1.06	−0.15	−0.56
Leverage	−0.37	0.52	1.02	−0.31	−1.15
Currency sensitivity	0.06	−0.05	0.26	0.20	0.75
Yield	0.25	−0.26	0.57	0.42	1.56
Non-EST universe	−0.16	0.23	1.77	−0.10	−0.38
Industries	0.56		4.44	0.18	0.66
Asset selection	5.09		5.24	0.97	3.63
Transaction cost	−2.46				

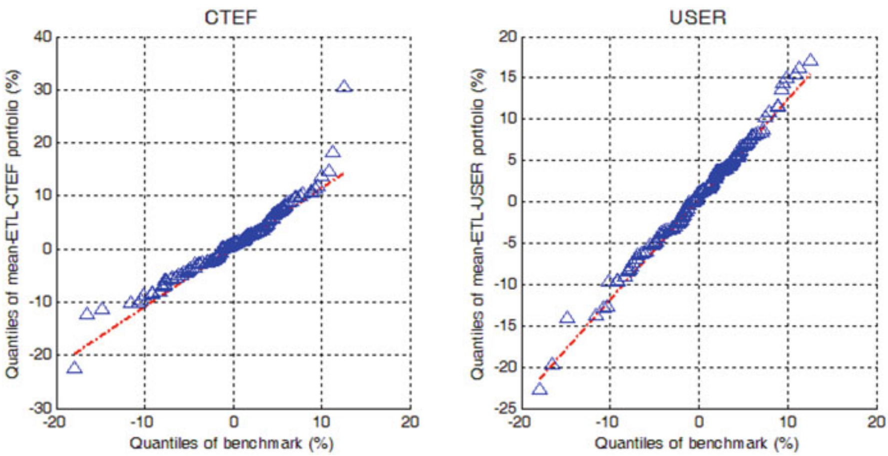


Fig. 7.1 Quantile–Quantile plot of portfolio monthly returns

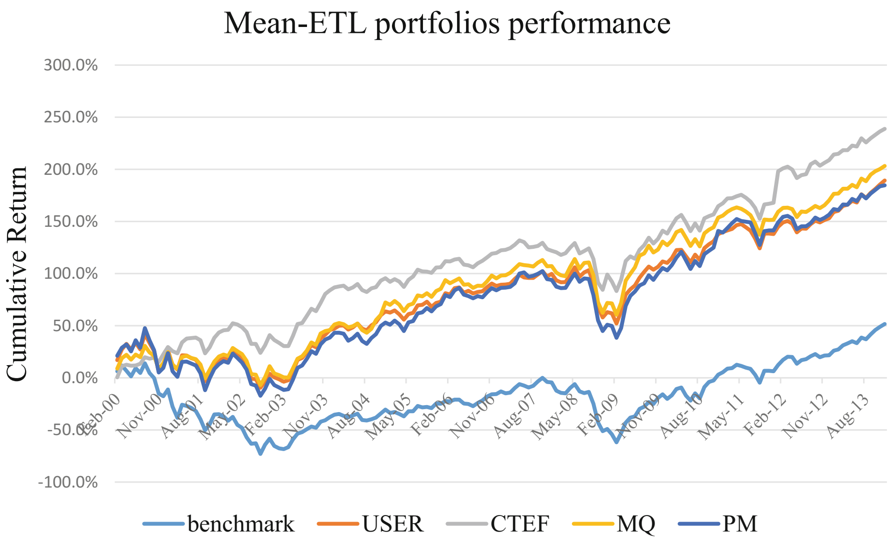


Fig. 7.2 Mean-ETL portfolios performance

Table 7.5 Performance criteria of mean-ETL created portfolios (US market), 2000–2013

	Mean-ETL portfolio				
	CTEF	**USER**	MQ	PM	Benchmark
Geometric mean	16.12 %	11.78 %	12.80 %	10.56 %	2.05 %
STD	20.87 %	21.84 %	22.86 %	25.03 %	18.30 %
Beta ($\boldsymbol{\beta}$)	0.92	1.10	1.10	1.26	
Sharpe ratio	0.68	0.45	0.47	0.34	
Treynor ratio	**0.15**	**0.09**	**0.10**	**0.07**	

characteristic lines of mean-ETL-CTEF and the benchmark index, showing that the mean-ETL-CTEF portfolio is outperforming the benchmark index (see Treynor & Mazuy 1966, for details).

We also find that the mean-ETL-MQ portfolio slightly outperforms the other portfolios except for the CTEF portfolio as shown in Table 7.5. As we described before, the MQ variable is a combination of CTEF and PM variables. The better performance of the mean-ETL-MQ portfolio is partly due to the strong performance of CTEF variable. The mean-ETL-PM portfolio has the lowest return and highest β, making it the worst performing portfolio in the four mean-ETL portfolios. Menchero (2015) also finds that momentum pure factor portfolio is more volatile than the earnings yield factor portfolio. Nevertheless, the mean-ETL portfolio based on PM variable still outperform the benchmark and produce statistically significant active returns as mentioned in the attribution report.

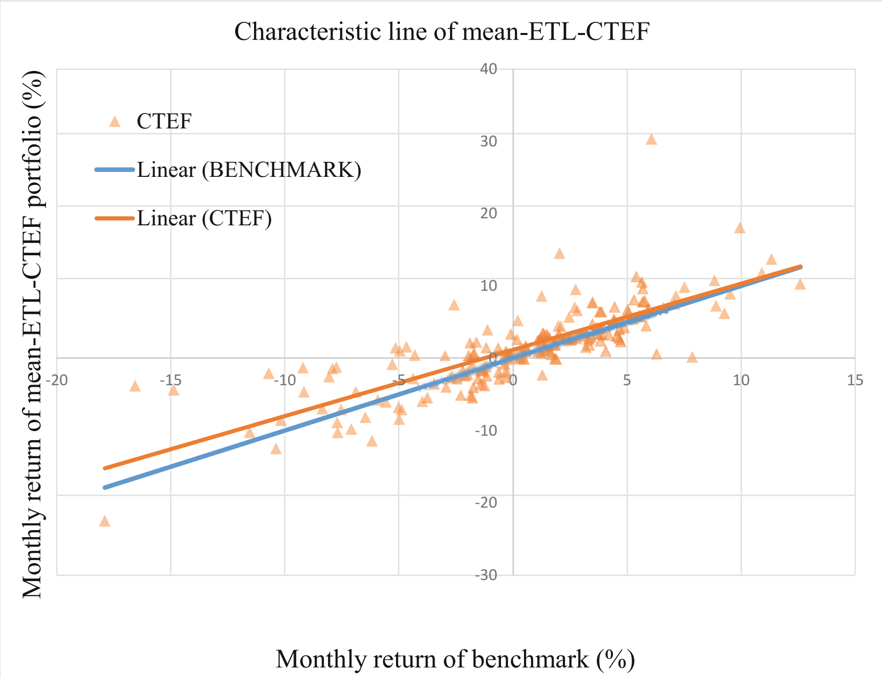


Fig. 7.3 Characteristic line of mean-ETL-CTEF portfolio

7.5 Summary

In this study, we focus on US stock markets and apply the mean-ETL portfolio construction on stocks' fundamental variables, CTEF, USER, PM, and MQ. We report that these fundamental variables continue to be very valuable for stock selection and portfolio construction. Also, the returns of the mean-ETL portfolios have thinner left tails, while not deteriorating the right tails. Mean-ETL portfolio with fundamental variables can generate statistically significant active returns in domestic market through asset selection, similar to its application to the global markets (see Guerard et al. 2013 & Shao et al. 2015). In this study, we also find that the mean-ETL-CTEF portfolio has the highest risk-adjusted return.

Disclosure The views and opinions expressed in this chapter are those of the author and do not represent or reflect those of Crabel Capital Management, LLC.

References

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance, 9*(3), 203–228.

Bloch, M., Guerard, J., Markowitz, H., Todd, P., & Xu, G. (1993). A comparison of some aspects of the US and Japanese equity markets. *Japan and the World Economy, 5*(1), 3–26.

Brown, L. D. (1999). Annotated I/B/E/S Bibliography.

Brush, J. (2001). price momentum: A twenty year research effort. *Columbine Newsletter.*

Chan, K., Hameed, A., & Tong, W. (2000). Profitability of momentum strategies in the international equity markets. *Journal of Financial and Quantitative Analysis, 35*(02), 153–172.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance, 47*(2), 427–465.

Graham, B., Dodd, D. L. F., & Cottle, S. (1934). *Security analysis* (pp. 44–45). New York: McGraw-Hill.

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American Economic Review, 1*, 1088–1105.

Guerard, J. B., Blin, J., & Bender, S. (1998). Forecasting earnings composite variables, financial anomalies, and efficient Japanese and US portfolios. *International Journal of Forecasting, 14*(2), 255–259.

Guerard, J. B., Gültekin, M., & Stone, B. K. (1997). The role of fundamental data and analysts' earnings breadth, forecasts, and revisions in the creation of efficient portfolios. *Research in Finance, 15*, 69–92.

Guerard, J. B., Rachev, S. T., & Shao, B. P. (2013). Efficient global portfolios: Big data and investment universes. *IBM Journal of Research and Development, 57*(5), 1–11.

Jaffe, J., Keim, D. B., & Westerfield, R. (1989). Earnings yields, market values, and stock returns. *The Journal of Finance, 44*(1), 135–148.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65–91.

Kim, Y. S., Giacometti, R., Rachev, S. T., Fabozzi, F. J., & Mignacca, D. (2012). Measuring financial risk and portfolio optimization with a non-Gaussian multivariate model. *Annals of Operations Research, 201*, 325–343.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.

Markowitz, H. (1959). *Portfolio selection: efficient diversification of investments*. : Yale University Press.

Menchero, J. (2015). *Performance of earnings yield and momentum factors in US and international equity markets*. This issue.

Rachev, S. T., Martin, R. D., Racheva, B., & Stoyanov, S. (2009). Stable ETL optimal portfolios and extreme risk management. *Risk Assessment Contributions to Economics, 2009*, 235–262.

Rachev, S. T., Stoyanov, S., & Fabozzi, F. J. (2007). *Advanced stochastic models, risk assessment, and portfolio optimization: the ideal risk, uncertainty, and performance measures*. New York: John Wiley & Sons.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk, 3*, 21–41.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance, 26*, 1443–1471.

Rosenberg, B., & Marathe, V. (1975). Tests of capital asset pricing hypotheses (No. 32). University of California at Berkeley.

Shao, B. P., & Rachev, S. T. (2013). Mean-ETL optimization of a global portfolio. *The Journal of Investing, 22*(4), 115–119.

Shao, B. P., Rachev, S. T., & Mu, Y. (2015). Applied mean-ETL optimization in using earnings forecasts. *International Journal of Forecasting, 31*(2), 561–567.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 1*, 119–138.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*(1), 63–75.
Treynor, J., & Mazuy, K. (1966). Can mutual funds outguess the market. *Harvard Business Review, 44*(4), 131–136.
Xia, H., Min, X., & Deng, S. (2015). Effectiveness of earnings forecasts in efficient global portfolio construction. *International Journal of Forecasting, 31*(2), 568–574.

Chapter 8
Portfolio Performance Assessment: Statistical Issues and Methods for Improvement

Bernell K. Stone

8.1 Introduction: Purposes and Overview

8.1.1 Performance Assessment Problems/Frameworks

Building on Markowitz (1952, 1959) mean–variance portfolio theory and the capital asset pricing model, Treynor (1965), Treynor and Mazuy (1966), Sharpe (1966), and Jensen (1968) set frameworks for portfolio performance assessment. Investment texts now all include chapters summarizing these measures. The crux of these performance assessment frameworks is assessing and explaining the amount that realized return exceeds a fair return for time and risk.

The primary focus of much performance measurement is after-the-fact assessment of how a managed portfolio performed relative to a before-tax fair return for time and risk. A related performance measurement problem is the task of evaluating well methods for active stock selection. The focus here is a full sample backtest of the performance potential of a stock return forecast. The assumed assessment structure is a panel framework for a time series of cross sections rank-ordered into fractile portfolios on the basis of a return forecast.

To assess with high statistical confidence the economic potential of a stock return forecast, the central backtest problem is to ensure that any apparent ability of a return forecast to predict future returns is well isolated from risk, tax, and other nonforecast return variables. The conventional methodology for correcting a cross section of realized returns for variation in risk is a multivariate regression using one

B.K. Stone (✉)
Global Energy Management, 4681 Windsor Drive
e-mail: BKS@BYU.EDU

© Springer International Publishing Switzerland 2017
J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_8

of the standard return models. Estimating how well-realized returns or realized risky returns are explained by a return forecast and any of the APT or multivariate style models such as the Fama–French three-variable extension of the CAPM is fraught with measurement and specification problems, especially extreme multicollinearity problems.

8.1.2 Purposes

This chapter presents and illustrates the use of an alternative return forecast assessment framework. Rather than estimating a multivariate explanation of how realized returns or realized risky returns depend on a return forecast and other explanatory variables, the proposed alternative suppresses cross-sectional variation in the other explanatory variables by transforming the initial cross section of rank-ordered fractile portfolios into an associated cross section in which every portfolio has the same portfolio-weighted average value of each pertinent explanatory variable.

Response surface/subsurface statistical designs are intended to assess the response (dependent variable) to a treatment. Response surface/subsurface designs are widely used in controlled experiments, e.g., chemical synthesis, petroleum refining, nuclear reaction yield, etc. In fact, response surface/subsurface designs are the preferred statistical design framework for most controlled experiments.[1] Response surface/subsurface designs are also widely used in partially controlled experiments such as the illustrative example in Sect. 8.4.2 of health response to well-controlled variation in drug dosage administered to a patient sample designed to be matched on other sample attributes that could distort apparent response to the drug dosages.

Response surface/subsurface statistical designs and methods are not widely used in economics and finance (or social sciences generally) although response surface/subsurface methods are implicit in the extensive use of matched-sample and partially-matched-sample designs, especially in areas like marketing research and medicine. Regression and related econometric methods are the generally preferred

[1]There is a large literature on statistical designs for the empirical estimation of multivariate functional dependencies but primarily focused on controlled or partially controlled studies rather than the observational samples that typically arise in epidemiology, demographics, the social sciences including especially economics and business, medicine, and many physical sciences including astronomy. Treatment response studies were an early statistical design focus and continue to be an ongoing design estimation concern. Because of the early and ongoing concern for treatment response studies, it is common to use the term **response surface methods** to refer to empirical methods for estimating functional dependencies. Most of the statistical literature pertains to controlled or partially controlled experiments in process industries, medicine, and advertising. For background readers are referred to the foundation works by Box (1954) and Box and Draper (1959); to the review article by Myers, Khuri, and Cornell (1989); and to the books on response surface methods: Khuri and Cornell (1996), Box and Draper (1987), and Box, Hunter, and Hunter (2005).

statistical method for assessing return dependencies in asset pricing studies and other return-risk modeling. In addition to goodness of fit measures such as standard forecast error or mean absolute forecast error, the standard procedure for evaluating a return forecast, i.e., for assessing forecast value, is to use regression to evaluate how well the cross-section of realized returns for a pertinent stock sample and time period is explained by the return forecast. As in the case of many return dependency studies, it is standard procedure to rank order the sample stocks on the predicted return or an adjusted return prediction measure such as return in excess of estimated risk. When the realized return response for the rank-ordered sample is estimated via linear regression, the estimated slope coefficient is referred to as the "*information coefficient*." Return forecast value and reliability is indicated by the magnitude and significance of the information coefficient.

The key conceptual insight is to view a stock return forecast as a treatment applied to all the stocks in a sample that is designed to rank order the stocks on the basis of true performance corrected for specified risk, tax, and other return impact variables. Rather than matched controls via sample selection, the control matching for an observation sample of stocks is achieved after the fact by means of a power optimizing mathematical assignment program that transforms the initial cross section of forecast rank-ordered portfolios into an associated rank-ordered cross section matched on key controls.[2]

In addition to presenting the matched control framework, the intent here is a systematic structuring of the major design decisions for a panel study based on a time series of rank-ordered fractile portfolios. Compared to forecast performance assessment using multivariate regression, the matched control framework has substantial efficiency/power benefits. Compared to the multicollinearity distortion in a multivariate regression, control matching on all the pertinent distortion variables means no correlation between the forecast and any of the return control variables. For instance, having the same portfolio average value of a variable such as beta means no cross-sectional variation in beta and therefore no correlation between beta and the return forecast or in fact any other control variable. Control matching ensures that the cross-sectional impact of the forecast is well isolated from any distortion from any of the control variables, because each portfolio in the cross-section has the same security-weighted average value of each impact variable. Since there is no variation in the portfolio-average value of each control variable over the

[2]Most of the statistical design literature cited in the previous footnote focuses on controlled and especially partially controlled studies. The ability to adapt response surface methods to observational studies is developed in Stone, Adolphson, and Miller (1993). They extend the use of response surface methods to observational data for those estimation situations in which it is pertinent to group data, e.g., to group observations on individual stocks into portfolios or households into income percentiles. The use of control variables to assess a conditional dependency (response subsurface) is a basic technique in controlled experiments that we adapt in this study to obtain a well-isolated portfolio-level dependency of realized risky return on a return forecast. Fortunately for compatibility with other finance panel studies using rank-ordered grouping, the optimal control matching can be structured as a power/efficiency improvement to the widely used relative rank-ordering used in many return dependency studies.

rank-ordered return cross section, there can be no variation over the cross section in the impact of any matched control to the extent that its average value is accurately measured and is a good summary of the impact of each control on the portfolio return.

Power pertains to correctly assessing sample information. In the context of a full-sample return forecast assessment, there are three pertinent power attributes. The most critical is isolation of the forecast from other return impact variables. The other two pertain to assessing well the magnitude and significance of the apparent forecast response. Control matching ensures complete isolation of the return forecast from the impact of other variables to the extent that control variables are well measured and well summarized by their portfolio average values. Assessing well the magnitude and significance of the forecast response is achieved via the use of a reassignment algorithm that transforms an initial forecast rank ordering into an associated control-matched rank ordering using the reassignment algorithm formulated in Section 8.5. This algorithm optimizes a trade-off between two complementary power instruments – having a wide range of well-ordered forecast values and also having a high level of variable homogeneity within each fractile portfolio. *Variable homogeneity* refers to the extent to which values are close to the portfolio average. Variable homogeneity is measured in this research study by the variance relative to the mean for each fractile portfolio.

8.1.3 Chapter Organization

The rest of this chapter is organized as follows. Section 8.2 distinguishes between standard forecast accuracy measures such as standard forecast error and the ability of a forecast to predict return beyond a fair return for time and risk. Section 8.3 focuses on key statistical designs that determine power and efficiency in estimating performance potential for a well-isolated return forecast. Section 8.4 compares multivariate regression with control matching as alternative ways to assess the performance potential of a stock return forecast. Section 8.5 formulates a mathematical program to transform the starting, presumably correlation-distorted, rank-ordered cross section into an associated cross section of well-ordered, control-matched portfolios with zero correlation with a specified set of control variables. The objective function is to optimize a trade-off between two power measures—cross-sectional range and within-portfolio variable homogeneity on the rank-ordering variable. The decision variable is the amount of stock in each fractile portfolio that is assigned to one or more of the other portfolios. The key constraints are the control matching requirement that each portfolio in the cross section has the same portfolio-weighted average value of the specified control variables.

Sections 8.6, 8.7, and 8.8 provide an illustrative performance potential assessment using an implementation of the eight-variable return forecast model of Bloch, Guerard, Markowitz, Todd, and Xu (1993). Section 8.6 provides an overview of the return forecast model. Section 8.7 defines and discusses a set of firm-specific control variables. With an emphasis on the elimination of risk and tax effects (both dividend–gain distortion and the tax shield of corporate debt), Sect. 8.8 illustrates the imposition of matched controls for different combinations of control variables. Section 8.9 summarizes conclusions and suggests issues for further research.

8.1.4 Overview of Some Key Results/Conclusions

In addition to optimizing statistical power, the imposition of a combination of risk and tax controls significantly increases statistical efficiency relative to the uncontrolled cross section. For the illustrative forecast model, tax effects are much larger than risk effects as measured by the three Fama–French risk variables: beta, size, and the book-to-market ratio. The three risk variables tend to smooth both the cross section of realized risky returns and especially the cross section of realized standard deviations. Both the return and the realized standard deviation cross sections are nonlinear. The cross section of realize standard deviations is not only nonlinear but highly nonmonotonic. As expected for a forecast designed to identify undervalued stocks (high upside potential with limited downside risk) versus overvalued stocks (limited upside potential with high downside risk), the distribution of realized returns about the average value exhibits significant skewness, negative skewness for low return forecasts, very little for the middle of the distribution, and very large significant skewness for the highest return forecasts.

Both the significant nonmonotonic cross section of realized standard deviations and the significant cross-sectional variation in realized skewness for the illustrative forecast attest to the importance of avoiding restrictive distributional assumptions. Overall, the matched control approach not only ensures a well-isolated return forecast but also provides significant improvements in both statistical efficiency and power compared to estimating a multivariate regression.

8.2 The Problem of Assessing the Performance Potential of a Stock Return Forecast

8.2.1 Forecast Accuracy/Significance Versus Performance Potential

Standard methodology for evaluating a forecast is to see how well actual values correspond to predicted values. For a stock return forecast, this realization versus forecast assessment translates into seeing how well-realized returns correspond to predicted returns. Standard forecast evaluation measures include the standard forecast error and the information coefficient. The standard procedure for computing an information coefficient is to estimate a linear fit of realized returns on predicted returns. The slope measures the information value of the forecast. A slope that is insignificantly different from zero implies no forecast value. Of course, a significantly negative slope implies negative forecast value. For a significantly positive slope, the larger the estimated slope and the higher the R^2 (the more significant the estimated slope), the better the forecast.

In commenting on a return forecast model of Timmerman (2008a), Brown (2008) asserted that high statistical significance for predicting future returns does not ensure actual ability of a stock return forecast to provide a significant improvement in performance after correcting returns for risk and other systematic return variables. In response, Timmerman (2008b) concurred with Brown's assertion that the appropriate forecast assessment criterion is the ability to create superior portfolio-level performance. However, Timmerman observed that such an assessment was itself problematic. In particular, converting realized return from a forecast into realized return corrected for time and risk means that the assessment is a joint test of the forecast and an assumed fair return model. In questioning the ability to provide a risk correction with a high degree of confidence, Timmerman (2008b) assumes that correcting realized return for risk requires a fair return model with its associated limitations. The issues raised by Timmerman are indicative of a need for a specification-free, distribution-free alternative methodology for obtaining the well-isolated performance potential assessment demanded by Brown (2008).

8.2.2 Key Specification Issue: Eliminating/Controlling for Correlation Distortion

The primary requirement for isolating well true return forecast value from risk, taxes, and other return dependency variables is to ensure that there is no significant distortion from covariation between the return forecast and these other variables. While rank-ordered grouping into fractile portfolios can increase statistical efficiency by mitigating measurement error as discussed further in Sect. 8.3.2, rank-ordered grouping can exacerbate the problem of eliminating correlation distortion. Low sample-level correlation between the return forecast and other return impact variables can be greatly multiplied by the rank-ordered grouping.[3] For a sample of 1000 stocks rank-ordered into deciles, relatively low sample-level correlation coefficients of 0.05 to 0.10 can be magnified to portfolio-level correlation coefficients greater than 0.50 and even as high as 0.80. Getting the measurement error reduction

[3]The mathematics for correlation magnification is straightforward. The formula for the correlation coefficient between variables X and Y is covariance(X, Y)/[SD(X)SD(Y)], where SD stands for standard deviation. Ranking on variable X and grouping into fractile portfolios preserve a wide range of values for variable X. However, in each of the portfolios, the individual values of variable Y tend to average out to a value close to the sample average value. Thus, the portfolio-level standard deviation of Y is reduced, and for a very small number of portfolios (e.g., quintiles or deciles), SD(Y) tends to approach zero while the covariance in the numerator declines relatively slowly because of the wide range of portfolio-level values for the ranking variable X.

benefits of rank-ordered grouping clearly requires a statistical design that explicitly deals with correlation distortion.

8.2.3 Eliminating/Controlling for Systematic Tax Effects: Dividends Versus Gains

Tax effects associated with cross-sectional variation in the proportion of return realized as dividends and capital gains[4] are a clear omission from all the standard before-tax APT and multivariate style models. A forecaster seeking to beat any of the before-tax return models can generate a rolling time series of forecasts that exploit the dividend yield tilt, the well-known difference in the before-tax return of dividends and gains reflecting the differential dividend–gain taxation and possible differences in the systematic risk of dividends and gains.

Trying to reflect tax effects by adding a dividend yield term or other dividend–gain mix explanatory term to a before-tax return model can further exacerbate covariation resolution issues. For instance, dividend yield is correlated or partially correlated with beta, size, and especially value variables such as the book–price ratio.

8.3 A Framework for Optimal Statistical Design

8.3.1 Key Design Decisions

Statistical efficiency and power are two complementary dimensions of how well a researcher can extract information from a data sample. Key attributes of information extraction include (1) sample size and how well sample information is used; (2) the usual estimation issues of measurement error, specification error, and omitted variable distortion; and (3) breakdown in the assumptions that underlie estimation–inference tests. Of particular concern in isolating return forecast performance from other return dependency variables is covariation contamination that can significantly distort efforts to isolate the impact of the return forecast from other variable dependencies.

[4]Brennen (1970) shows that dividend yield is a significant omitted variable from the CAPM. Rosenberg (1974), Rosenberg and Rudd (1977), Rosenberg and Marathe (1979), Blume (1980), Rosenberg and Rudd (1982) and many subsequent researchers have empirically established the so-called dividend yield tilt. More recent studies include Peterson, Peterson, and Ang (1985), Fama and French (1988) and Pilotte (2003). For an extensive review of both dividend valuation and dividend policy and extensive references in this area, see Lease, Kose, Kalay, Loewenstein, and Sarig (2000).

Given the decisions on the forecast, the pertinent stock sample, and the time frame for the backtest assessment, the major power–efficiency decision is the methodology used to isolate the return forecast from other return impact variables. As already indicated, the conventional methodology is a multivariate regression. The alternative methodology advocated here is the use of matched controls.

8.3.2 The Number of Fractile Portfolios: Measurement Error Versus Power

Two key decisions for rank-ordering into fractile portfolios are the number of fractile portfolios in a cross section and the partitioning rule for deciding on the number of securities in each portfolio. These two decisions are critical for statistical efficiency related to measurement error versus three statistical power attributes: (1) portfolio sample size, (2) cross-sectional forecast range, and (3) within-portfolio variable homogeneity. *Within-portfolio variable homogeneity* refers to how well the portfolio average value represents the collection of stocks within the portfolio. Fewer fractile portfolios and therefore more stocks per portfolio mean that return dependency variables including the rank-ordering return forecast will generally have greater dispersion about the portfolio average value. Within-portfolio variance is a measure of the departure from the mean.[5]

In addition to reducing nonsystematic return variation, the primary reason for grouping stocks into fractile portfolios is to reduce measurement error. Early rank-ordered grouping paradigms established in tests of the capital asset pricing model in Fama and MacBeth (1973) and in Black, Jensen, and Scholes (1972) used rank-ordered grouping into deciles to mitigate beta measurement error. Individual stock betas have relatively large estimation errors and tend to change over time. To the extent that beta measurement errors are independent of each other,[6] estimation measurement error for a portfolio of 50 stocks is reduced to approximately 1/50 of the average stock-level measurement error.

[5]Other dispersion measures could be used, e.g., mean absolute deviation or interquartile range. Relative to these measures, within-portfolio variance gives greater weight to extreme departures from the portfolio average.

[6]Since beta measurement errors are known to regress toward the mean, the assumption of uncorrelated measurement errors used in the discussion here is almost certainly too strong for correctly assessing efficiency gains from reducing measurement error by grouping into rank-ordered fractile portfolios. In particular, beta values that are correlated with the forecast are very likely to have systematic variation in beta change values over the cross section. The control matching methodology developed in Sects. 8.4 and 8.5 will mitigate systematic changes in measurement error such as the well-known regression of betas toward the mean. When every portfolio in the cross section has the same beta value, each portfolio will have essentially the same ex post beta change. Hence, control matching provides the benefit of no systematic distortion from beta regression toward the mean.

In designing a study to evaluate forecast value, there is a clear need to trade-off the efficiency gains from measurement error reduction (see for instance Grunfeld and Griliches (1960) and Griliches (1986) for background) versus the information/power loss from reduced sample size and the associated power costs of reduced cross-sectional range (cross-sectional variance) and loss of within-portfolio variable homogeneity.

To illustrate the efficiency/power trade-off, consider partitioning a sample of 1000 stocks into equal-size portfolios of 200, 100, 50, and 20 stocks per portfolio. This partitioning results in 5, 10, 20, and 50 portfolios, respectively. Given that a portfolio of 50 and 20 stocks provides 98 % and 95 % of the measurement error reduction for uncorrelated measurement errors, collapsing the sample size to just ten portfolios (deciles) or even more extremely just five portfolios (quintiles) seems to be a clear case of excessive measurement error reduction relative to lost power from collapse of the sample size.

8.4 Isolation Methodology Alternatives: Multivariate Regression Versus Control Matching

8.4.1 Treatment Response Studies

To motivate intuition, it is useful to think of assessing stock return forecasting performance potential within the broad class of treatment response studies. A stock return forecast is a treatment applied to a sample of stocks to identify misvalued stocks. The assessment problem is to be sure that any apparent ability to rank order a stock sample on the basis of superior return is from the ability to separate true misvaluation from risk, taxes, and possibly other nonforecast return responses.

A *treatment response assessment* is a special case of empirically estimating a functional dependency.[7] The response is assumed to depend on the treatment and other explanatory variables. However, the usual concern in a treatment response assessment is not necessarily estimating the overall response dependency on the

[7]There is a large literature on statistical designs for the empirical estimation of multivariate functional dependencies but primarily focused on controlled or partially controlled studies rather than the observational samples that typically arise in epidemiology, demographics, the social sciences including especially economics and business, medicine, and many physical sciences including astronomy. Treatment response studies were an early statistical design focus and continue to be an ongoing design estimation concern. Because of the early and ongoing concern for treatment response studies, it is common to use the term *response surface methods* to refer to empirical methods for estimating functional dependencies. Most of the statistical literature pertains to controlled or partially controlled experiments in process industries, medicine, and advertising. For background readers are referred to the foundation works by Box (1954) and Box and Draper (1959); to the review article by Myers, Khuri, and Cornell (1989) and to the books on response surface methods: Khuri and Cornell (1996), Box and Draper (1987), and Box et al. (2005).

treatment and all other explanatory variables but rather ensuring that the estimated response to the treatment is well isolated from distortion from other explanatory variables. Using regression to estimate a multivariate response dependency is one way to assess treatment response. However, a good statistical design can significantly increase both efficiency and power by creating a series of treatment observations that are matched on the values of one or more of the nontreatment explanatory variables. The matched variables are called *control variables*, often shortened to *controls*. Matched controls eliminate treatment variation from all the controls.

While purely controlled experiments such as petroleum blending–refining and chemical synthesis are a widely studied class of treatment response statistical designs, partially controlled studies such as drug response and product market response to pricing are more pertinent analogs for assessing how realized risky return responds to a stock return forecast.[8]

8.4.2 Intuition Motivation: Isolating Well Treatment Response to Drug Dosage Variation

A test of an anti-inflammatory drug could look at dosage per unit of body weight to assess inflammation reduction. The prototypical treatment response study is conducted on double-blind subsamples, with treatments ranging from no drug (the placebo) and then a range of well-structured dosage increases up to a maximum safe treatment. Each subsample is selected to match each other as much as possible on control variables such as sex, age distribution, obesity distribution, blood pressure, etc. The term "control variable" here refers to attributes of the study population that can impact initial inflammation and inflammation response or otherwise distort assessment of response to the drug dosage. The creation of the subsamples from an initial study population is designed to provide identical values on the key control variables. If all subsamples are matched on all pertinent controls, then inflammation-dosage response is well isolated from variation in the control variables.

Because of initial subsample differences and especially because of withdrawals and disqualification of some of the treatment subjects, the cross section of efficacy assessment subsamples is usually not a perfect match on all controls. Researchers

[8]Most of the statistical design literature cited in the previous footnote focuses on controlled and especially partially controlled studies. The ability to adapt response surface methods to observational studies is developed in Stone et al. (1993). They extend the use of response surface methods to observational data for those estimation situations in which it is pertinent to group data, e.g., to group observations on individual stocks into portfolios or households into income percentiles. The use of control variables to assess a conditional dependency (*response subsurface*) is a basic technique in controlled experiments that we adapt in this study to obtain a well-isolated portfolio-level dependency of realized risky return on a return forecast.

assessing dosage response must deal with possible distortion from variation in the unmatched controls.

What is pertinent for intuition about statistical procedure for cross-sectional assessments of portfolio performance potential are alternative ways that researchers can deal with cross-sectional variation in unmatched controls to ensure that drug efficacy assessment is well isolated from distortion in unmatched controls. One alternative is to use a multivariate regression model that attempts to explain observed inflammation changes by a combination of dosage variation and a regression-estimated inflammation response to other variables. Given the very difficult problem of modeling inflammation response to other variables and given the generally small departure from a match on the intended controls, the preferred alternative to a multivariate regression assessment is a transformation of the unmatched cross section back into a new cross section of control-matched subsamples. In drug dosage studies, the sample transformation may be accomplished by holding out some subsample observations. For instance, subsamples with above- and below-average obesity values can remove some above-average and below-average obesity observations from the respective subsamples.

Rather than giving some sample observations a de facto weight of zero, the rematching problem is usually structured as an optimization problem in which observations in each subsample are reweighted. The objective is to find the overall reweighting that minimizes the reduction in sample size (measured as the overall departure of all observation weights from one) while producing a sufficiently near match on each pertinent control variable. The benefit of the control rematching approach is that drug efficacy assessment is reduced to the intended evaluation of a univariate response to the treatment differences. The difficult multivariate regression estimation problem with its associated limitations in functional form modeling has been bypassed. **Rather than spreading the sample observations over the estimation of a multivariate dependency, all of the sample data can be concentrated on the univariate response to the varying dosage treatments**. The statistical properties (explained variation, t-value on treatment, F-stat) of the univariate response almost always dominate the corresponding statistical measures for the multivariate estimation.

The key point of this rather long discussion of a prototypical drug treatment response study is the significant efficiency/power benefits associated with the use of matched controls. The concern in a treatment response assessment is response to the treatment, a conditional univariate dependency. Variation in the controls is a source of distortion in assessing the treatment response. Trying to estimate the control impact involves the unnecessary use of sample data to eliminate distortion from control variable variation. Producing a match on each control avoids this very difficult and generally unnecessary estimation problem and focuses all sample information on the conditional univariate dependency of concern.

In viewing a forecast as a treatment applied to the stock sample in which we want to observe the performance response, the concern is again the estimation of a conditional univariate dependency. The impact of cross-sectional variation in firm-specific values of risk variables, tax effect variables, and possibly other firm impacts such as differences in growth and profitability can be eliminated from the cross section by transforming the initial rank ordering into an associated cross section that is matched on key return–risk impact variables.

Control matching is especially pertinent to backtest the return response to a forecast using time series of rank-ordered portfolios because grouping stocks into fractile portfolios can magnify sample-level correlation distortion. The magnification is nonlinear so that very low stock-level correlations can be multiplied dramatically when the overall sample is collapsed to a very small number of portfolios. The control-matched cross section has no cross-sectional variation in the control variables and thus zero correlation with the rank-ordering variable.

8.4.3 Transforming a Rank-Ordered Cross Section into a Control-Matched Cross Section

Ranking on forecasted return and grouping into fractile portfolios will produce a set of portfolios ordered on the basis of predicted return. This return cross section will almost certainly have a wide range of forecasted return values. However, each portfolio in the cross section will almost never have the same average values of explanatory variables such as beta or size or the dividend–gain mix or any of the other return impact variables listed in Exhibit 8.1.[9] To the extent values of return impact variables fluctuate randomly about their average value over the cross section, their variation is primarily a source of noise and therefore a source of lost efficiency in assessing the cross-sectional dependency of realized returns on the return forecast score.

Much worse than lost efficiency from random variation in systematic risk, tax, and other return impact variables is the problem of distortion from correlation or partial correlation between these return dependency variables and the return

[9]Candidates for control variables are any variable believed to have a significant impact for explaining or predicting the cross-section of realized returns. Classes of return impacts include risk measures, e.g., beta, the book-price ratio, and firm size; tax valuation impacts, e.g., dividend yield, the dividend payout ratio, and possibly the debt tax shield as measured by the percentage of financing that is debt; and attractiveness measures that are indicative of future cashflow generation potential and asset usage efficiency, e.g., growth, return on investment, or sales intensity (sales per dollar of investment). Beta is the standard measure of volatility risk established as a return explanatory variable in formulations of the capital asset pricing model, for instance Sharpe (1964). It is also included in multifactor return modeling, as indicated by the Fama-French series, e.g., Fama and French (1992, 1996 2008a, 2008b). The variables EP and BP are the reciprocals of the price-earnings ratio and the price-book ratio, respectively. Their use as valuation and/or risk variables has been researched extensively beginning with Basu (1977), viewing dependency on the earnings-price ratio primarily as a valuation anomaly but recognizing the possibility that the earnings-price ratio could also be a risk instrument. The tax effect associated with the differential taxation of dividends and capital gains and the debt tax shield are discussed extensively in Sect. 8.7.3.

VARIABLE NAME	SYMBOL	VARIABLE DEFINITION
Beta	β	$\beta = Cov(R_s - R_o, R_M - R_o) / Var(R_M - R_o)$ measured over 3 years of past monthly returns, where R_o is the riskless rate and R_M is return on a market index.
Book-to-Price Ratio (Book-to-Market Ratio)	**BP**	Ratio of book value to market value. Book value **BP** is the latest (at least two-month back) annual CRSP value for ***total common equity.*** Market value **MV** is the current market value of common stock (current stock price times number shares outstanding).
Size (Market Value of Common Stock)	**S**	The market value of common stock at a point in time
Earnings-Price Ratio (Earnings Yield)	**EP**	The ratio of **Net Income** to **market value**, the reciprocal of the price-earnings ratio
Dividend-Price Ratio (Dividend Yield)	**DP**	The ratio of **Annual Dividends** to **Market Value**
Financial Structure	**FL**	The fraction of **Total Investment** provided by debt and preferred stock
Sales Growth	**SAG**	Five-year average sales growth
Sustainable Growth	**SUG**	The growth of common equity from retained earnings
Return on Investment	**ROI**	The ratio of **Operating Income** (before extraordinary income and expenses) to **Total Investment**
Return on Equity	**ROE**	The ratio of **Net Income** to **Book Value**
Sales Intensity	**SI**	The ratio of Sales to **Total Investment**

Exhibit 8.1 Summary of control variables

forecast. To the extent that a systematic risk or tax variable is correlated with the return forecast score, the cross-sectional dependence of realized risky return will reflect the well-ordered cross-sectional change in the correlated variable. For instance, if a return forecast were positively correlated with beta, the observed cross section of realized risky returns will include the systematic variation in beta. An apparent increase in realized risky return from the return forecast will also include any return to beta risk bearing. Similarly, if the dividend–gain mix increases systematically with the return forecast, once again, an apparent increase in realized risky return with the return forecast can be a tax tilt in disguise.

The conventional methodology for separating return forecast potential from dependence on other variables is multivariate regression. For instance, if the concern were just to correct for beta risk in the context of the CAPM, a regression that adds the return forecast to the linear market index model as an additional explanatory variable would resolve the relative importance of beta and the return forecast in explaining the cross section of realized risky returns. For a multivariate risk correction, adding the return forecast to the three-variable Fama–French return model might remove the effect of the three risk variables on the cross section of realized risky returns but clearly incur a multicollinearity problem if the concern is correlation distortion from the Fama–French risk variables. Trying to correct as well for systematic tax effects is even more problematic. For instance, adding a term based on dividend yield to reflect variation in the dividend–gain mix adds another explanatory variable that is correlated with all the Fama–French variables, especially the book–price ratio.[10] Adding more variables to model better the cross section of realized risky returns, for instance, combining the return forecast with all three Fama–French risk variables plus the three tax variables (DP, EP, and FL in Exhibit 8.1), and then additional variables to reflect differences in growth and profitability means using up degrees of freedom while incurring more measurement error and creating an ever worse multicollinearity problem.

As in the previously discussed example of the drug treatment response assessment, isolating well-realized return response to a forecast does not require empirically measuring an overall return dependency but rather ensuring a well-measured return-to-forecast response. In the case of a stock return forecast, the primary isolation requirement is to eliminate distortion from correlated variables in the context of a good assessment design that appropriately trades off efficiency and power.

As in the drug treatment response assessment, the contention here is that the use of matched controls is superior to using a multicollinearity-contaminated multivariate regression when the goal is isolating the impact of a return forecast from other return impact variables. The isolation alternative is to transform the initially rank-ordered cross section into an associated cross section matched on the variables that would be used as explanatory variables in a regression. A control-matched variable has the same impact on each portfolio in the cross section. There is no cross-sectional variation from the matched control and therefore zero correlation distortion.

The drug treatment example was a partially controlled experiment. Each dosage subsamples was selected to match every other treatment subsample on key controls. Rematching to reflect withdrawals and excluded subjects may be accomplished by pruning and more generally by an optimal reweighting of the subjects in each subsample. In the drug dosage study, each subsample represents a well-ordered dosage change. Dosage observations in one subsample would not be mixed with dosage observations from another subsample. In contrast, each fractile portfolio has a distribution of return forecasts summarized by their average. Reassigning stocks is

[10]Trying to put before-tax returns on an after-tax basis is fraught with problems. To put the dividend component of return on an after-tax basis requires an estimate of a time-varying marginal tax rate for ordinary income. To put the gain component of return on an after-tax basis requires the determination of the time-varying effective tax rate on capital gains.

an alternative way to obtain a match on key explanatory variables. Thus, rather than either excluding some stocks having extreme values of some control variables[11] or the more general use of portfolio-by portfolio reweighting to produce a transformed cross section matched on key control variables, the approach developed here is a reweighting that allows stocks to be reassigned to adjacent portfolios.

8.5 A Power Optimizing Mathematical Assignment Program

As an example of cross-portfolio reassignment, assume trying to make each portfolio in the cross section have the same average beta value. Cross-portfolio reassignment could move a stock with an above-average beta value into a portfolio whose average beta value is below the population average. At the same time, a stock with a below-average beta value in a below-average beta portfolio could be shifted into an above-average portfolio.

Just to produce a match for each portfolio in the cross section on a single explanatory control variable such as beta clearly is computationally complex for a large stock sample with many fractile portfolios. There is a need for an objective algorithmic procedure to produce the best control-matched transformation. The problem of transforming an initial rank-ordered cross section into the best control-matched cross section can be formulated as a mathematical assignment program. Given an initial rank ordering, the criterion for "best control match" is to optimize three power measures. Covariation distortion is suppressed completely by the control matching while optimizing a trade-off between range and within-portfolio forecast variance.

8.5.1 Overview: Formulating the Mathematical Assignment Program

The assumed input to the control matching algorithm is a rank-ordered grouping into fractile portfolios. In addition to input data, an optimization requires specification of decision variables, an objective function, and constraints.

Given a cross section of fractile portfolios formed by rank ordering on predicted return, the objective of the assignment program is to transform this cross section into an associated control-matched cross section to optimize two complementary attributes of statistical power:

[11] In a stock return forecast designed to find misvalued stocks, extreme values of some return variables are very likely the observations of greatest performance potential.

1. Preserving a wide range of well-ordered return forecasts
2. Preserving within-portfolio homogeneity of forecasted return.

The following is a verbal statement of four generic constraints.

1. *Control match restriction*. For each fractile portfolio, make the portfolio average value of each control variable equal to the mean (average) value of that control variable in the sample population.
2. *Preserving initial portfolio size*. In reassigning securities to create the control matching, keep the number of securities in each of the fractile portfolios the same as the number of securities in that portfolio in the initial (starting) rank-ordered cross section.
3. *Full assignment*. Each security must be fully assigned.
4. *No short sales*. There can be no short sales.

The crucial constraints are the control matching restrictions. Preserving initial portfolio size and full use of each security are technical constraints that go with full use of the sample. Prohibiting short sales prevents one return observation from canceling out other return observations. Prohibiting short sales is also consistent with the idea of full use of all sample information in a long-only framework.

8.5.2 Notation Summary

The following summarizes notation and defines key variables.

$P =$ number of rank-based portfolios in the cross section
$p = 1$ is the portfolio with the smallest value of the rank-ordering variable
$p = P$ is the portfolio with the largest value of the rank-ordering variable
$S =$ total number of securities being assigned to portfolios
$s =$ security subscript
$\mathrm{FS}_s =$ the forecast score for stock s, $s = 1, \ldots, S$
$X_{ps} =$ the fraction of security s assigned to fractile portfolio p, $0 \leq X_{ps} \leq 1$
$F_p =$ the number of securities in fractile p in the starting rank ordering
$C =$ number of control variables
$V_c =$ control variable c, $c = 1, \ldots, C$
$\mathrm{VTARGET}_c =$ the target value for control variable c, $c = 1, \ldots, C$[12]
$D_{ps} =$ a difference measure of the change in rank for stock s when reassigned to portfolio p

[12] In this study, the target average value is always the ex ante sample average value.

8.5.3 *The Power Optimizing Objective Function*

Preserving range and minimizing cross-portfolio mixing are two aspects of statistical power. They are complementary measures in that optimizing one tends to optimize the other. To reflect the relative importance of these two measures, let Φ be a trade-off parameter that defines a relative weighting for range and within-portfolio variance, where $0 < \Phi < 1$. The trade-off between range and within-portfolio variance can be written as

$$\text{OBJECTIVE} = \Phi\,[\text{RANGE MEASURE}] - (1 - \Phi)\,[\text{Within-Portfolio Variances}] \tag{8.1}$$

For each portfolio in the cross section, the within-portfolio variance is the portfolio-weighted squared deviation of return forecast score from the portfolio mean forecast return score. It is a quadratic function. Thus, minimizing within-portfolio variance, actually minimizing a sum of within-portfolio variances over the cross section, means a quadratic objective function.

In this study in which we assess month-to-month return cross sections in each of the 456 months of 1967–2004, we impose progressively more complete sets of control variables in each month. Obtaining 15 or more control-matched cross sections in 456 months means solving more than 6700 optimization runs. Solving this many quadratic programs would be a computational challenge. However, just as one can approximate well the mean–variance portfolio optimization of Markowitz (1952, 1959) by solving an associated linear programming (LP) approximation to the quadratic program,[13] one can approximate the control matching quadratic optimization by an associated LP objective function.

The LP approximation objective function is

$$\text{Maximize}: \quad \text{LP OBJECTIVE} = \Phi\,[\text{RANGE}] - (1 - \Phi)\,[\text{SHIFTING}] \tag{8.2}$$

The linear measure SHIFTING is the approximation to variance minimization that we now define.[14] Let D_{ps} be the squared difference in the numerical rank between portfolio p and the natural portfolio rank of security s in the initial rank-order partitioning into fractile portfolios. The set of D_{ps} can be summarized by a symmetric PxS matrix. Squaring the difference means that all values are greater than zero. Squaring the difference also means that large shifts are much worse than small ones. If a stock stays in the initial portfolio, D_{pp} is zero for no shifting. If all or

[13] See, for instance, Sharpe (1963, 1967, 1971) and Stone (1973).

[14] It is intuitive that minimizing the amount and distance of cross-portfolio shifting tends to preserve the original within-portfolio forecast distribution including within-portfolio variances. The substance of this approximation is to use portfolio rank-order distance as a substitute for actual return forecast differences. Since we map each return forecast into a near uniform distribution on the (0, 1) interval, we tend to ensure the validity of this approximation.

part of a stock is shifted up or down by one, two, and three portfolios, the respective values of D_{ps} are 1, 4, and 8. Thus, reassignments of two or more portfolios up or down the rank ordering are highly penalized.[15]

If FS_s denotes the value of the forecast score for stock s, then the linear approximation objective function above can be written in terms of assignment variables as

$$\text{Maximize} \qquad \Phi\left[\Sigma_s X_{Ps} FS_s - \Sigma_s X_{1s} FS_s\right] - (1-\Phi)\left[\Sigma_p \Sigma_s X_{ps} D_{ps}\right] \qquad (8.3)$$

The mathematical assignment program can be solved for a range of trade-off values by varying Φ from zero to 1. In the results reported in Sect. 8.8, the value of the trade-off parameter Φ is 0.25. However, experience shows that the solutions are robust to variation in Φ. The reason for the robustness is that these two attributes of statistical power are complementary objectives. Minimizing cross-fractile shifting generally preserves most of the range as well as the distribution of return forecast scores in the starting fractile portfolios.

8.5.4 *Control Matching: The Equal Value Constraint for Each Control Variable*

Let V_s denote the security s value of a representative control variable. Let VTARGET denote the target value of this representative control variable for all P portfolios in the cross section. The representative control constraint can be expressed as

$$\sum\nolimits_s X_{ps} V_s = \text{VTARGET} \qquad p = 1, \ldots, P \quad \text{and every control variable} \qquad (8.4)$$

8.5.5 *Security Usage and Short Sales: Technical Constraints*

We impose two generic data usage constraints. The first says that each security must be fully assigned to one or more portfolios, i.e.,

$$\sum\nolimits_p X_{ps} = 1 \qquad s = 1, \ldots, S \qquad (8.5)$$

[15]The changed difference in changed rank is actually a stronger restriction on changing portfolio membership than the quadratic variance change it is approximating. Because the LP shifting measure penalizes very large rank shifts even more than the quadratic, the LP approximation tends to preclude large shifts in rank order even more than the quadratic. However, comparison of the LP and quadratic solutions showed that the LP and quadratic solutions were generally close.

The second security assignment constraint keeps the number of securities in each matched portfolio the same as the number of securities in the corresponding fractile of the starting rank-order partitioning of the distribution of $V1$. Let F_p denote the number of securities in fractile p. Then this restriction is

$$\sum_s X_{ps} = \mathbf{F}_{\mathrm{p}} \qquad p = 1, \ldots, P \tag{8.6}$$

The no short-sale restriction and the natural limitation that no security can be used more than once require

$$0 \leq \mathrm{X}_{ps} \leq 1 \qquad s = 1, \ldots, S \quad \text{and} \quad p = 1, \ldots, P \tag{8.7}$$

8.5.6 Synthesis of the Power Optimizing Reassignment Program

Optimization arises in finding the particular reassignment that optimizes a trade-off between preserving the widest possible range of well-ordered portfolio values of forecasted return and also ensuring preservation of within-portfolio homogeneity of forecasted return.

Given the sample of stocks with variable values for each stock in that time period, once we pick a number of portfolios P in the cross section and select a set of control variables, the transformation of the rank-ordered cross section into the control-matched cross section is defined by the optimization program. The mapping from the rank-ordered input cross section into the control-matched output cross section is objective in the sense that the forecaster/researcher exercises no discretion in how stocks are reassigned. The input cross section and the mathematical program determine the output cross section.

The substance of the reassignment process is well understood by knowing input and output. The input is a cross section of fractile portfolios. The rank-ordering variable is the return forecast. The overall output is a cross section of fractile portfolios that are matched on a specified set of controls variables. The mathematical program finds an optimal reassignment of stocks that transforms the input rank-ordered cross section into a new cross section that is matched on the portfolio average values of each control variable.

The input values of the assignment variables are the relative weighting of each stock in each portfolio in the cross section without any controls. The output values of the assignment variables are the relative weighting of each stock in each portfolio in the control-matched cross section.

The relative amount of each stock in each portfolio can be used to compute portfolio average values of pertinent portfolio attributes. Of most concern is the realized risky return for each portfolio. Given a time series of rank-ordered input portfolios and a corresponding time series of control-matched output portfolios, it is straightforward to compute the average realized risky return for each portfolio before and after controls and then to assess differences associated with the control variables.

8.6 Forecast Model Overview

8.6.1 Selecting an Illustrative Forecast Model

To illustrate well the benefits of using controls to isolate forecast performance from risk, tax, and other nonforecast impacts, a good illustrative forecast model should have a statistically significant dependency of realized risky returns on the return forecast. From the point of view of an illustration, it does not matter if the apparent dependency of realized returns on the stock return forecast is true alpha performance or is from risk, tax distortion, or other nonmodel return impact variables. In fact, when it comes to illustrating forecast isolation methodology, it is actually good if the cross-sectional return dependency is a mixture of effects from the return forecast itself and from systematic risk variables, tax effects, and other nonmodel return performance variables. In effect, to illustrate isolation methodology, it is actually good to have a "dirty return dependency" in the sense that the return dependency includes apparent performance from variables other than the forecast model itself.

The model selected to illustrate the benefits of the control methodology is an eight-variable, fundamental value-focused, rolling horizon return forecast model first published in Bloch, Guerard, Markowitz, Todd, and Xu (1993). We hereafter refer to this return forecast model as the BGMTX return forecast model. In talking about the generic approach of using a regression-estimated weighting of their eight value ratios, we shall refer to the BGMTX forecast approach or BGMTX forecast framework.

In addition to a very rigorous implementation in terms of only using data publicly available well ahead of forming the forecast, BGMTX assessed performance potential by using the model return forecast as the return input for a mean–variance portfolio optimizer. The other inputs to the mean–variance optimization were rolling horizon forecasts of security-level risk parameters.[16] The mean–variance optimizer transformed the rolling horizon return and risk forecasts into a time series of predicted mean–variance efficient portfolios in both Japan (first section, nonfinancial Tokyo Stock Exchange common stocks, January 1975 to December 1990) and the United States (the 1000 largest market-capitalized common stocks, November 1975 to December 1990). BGMTX reports that the mean–variance optimized portfolios

[16]For details on the mean–variance optimization used, see Markowitz (1959, 1987).

significantly outperformed benchmark indices even after testing for both survivor and backtest bias.[17]

8.6.2 Overview of the Illustrative Eight-Variable Forecast Model[18]

The BGMTX return forecast model uses a weighted average of eight value ratios[19]:

EP = [earnings per share] / [price per share] = earnings-price ratio
BP = [book value per share] / [price per share] = book-price ratio
CP = [cash flow per share] / [price per share] = cash-price ratio
SP = [net sales per share] / [price per share] = sales-price ratio.
REP = relative earnings-price ratio = EP/ [most recent five-year average value]
RBP = relative earnings-price ratio = BP/ [most recent five-year average value]
RCP = relative earnings-price ratio = CP/ [most recent five-year average value]
RSP = relative earnings-price ratio = SP/ [most recent five-year average value]

The first four ratios are called *current value ratios* in a sense of being the most recently reported values relative to the current price per share. Current value ratios measure value in terms of attractiveness compared to other peer companies. For instance, all other things being equal, a relatively high EP or BP ratio for a stock means that the stock is relatively more value attractive than the peer stocks with lower values for their EP and/or BP ratios.

The last four ratios defined above are *relative value ratios*. The "most recent five-year average value" in the denominator of these four relative value ratios means the five-year average of the ratio in the numerator. The four relative value ratios each

---

[17] Markowitz and Xu (1994) later published the data mining test for backtest bias. Their test allows assessment of the expected difference between the best test model and an average of simulated policies.

[18] BGMTX is a one-step direct forecast of stock returns. The more common return forecast framework is a two-step return forecast in which an analyst predicts both a future value of a variable such as earnings and an associated future value multiple for that variable such as a future price–earnings ratio. These two predictions imply a prediction of future value. Under the assumption that the current price will converge toward this predicted future value, there is an implied prediction of a gain return. Given a prediction of future dividends, there is an implied stock return forecast. For a thorough treatment of the two-step framework and extensive references to the two-step return prediction literature, readers are referred to the CFA study guide by Stowe et al. (2007). Because BGMTX is a direct one-step return prediction following a step-by-step determination of a normalized weighting of current and relative value ratios, it is amenable to a repeatable backtest.

[19] The ratio BP is of course the book-to-market ratio. BP is defined here as the ratio of book value per share to price per share. However, multiplying both numerator and denominator by the number of outstanding shares gives the ratio of book value to market value.

indicates relative attractiveness compared to a company's own past values of the four value ratios in the numerator. Thus, a stock is viewed as attractive not only when it provides a relatively higher earnings' yield than peer companies but also when it provides a high earnings' yield relative to its own past values. If a stock has a high relative EP ratio relative to the stock of peer companies, then that stock has had a greater relative decline in its price–earnings ratio and is thus a relatively "out-of-favor" stock.

These two types of value ratios arise from two complementary ways that fundamental value managers say they use value ratios, namely, (1) attractiveness relative to peer companies and (2) attractiveness relative to a company's own past valuations. In this sense, the relative weighting of these eight value variables can be thought of as a regression-based simulation of the type of fundamental value analysis advocated in works such as Graham and Dodd (1934) and Williams (1938).[20]

8.6.3 Variable Weighting: A Step-By-Step Implementation Summary

Having identified eight ratio variables as potential return predictors, the forecast modeling question is how to use these variables to predict future returns. An obvious way to evaluate relative predictive value is to assess how well they explain recent past returns. BGMTX uses regression to estimate the relative ability of these eight variables to explain past returns. Let R_s denote the return on stock *s* in a sample of *S* stocks. A linear regression equation to assess the relative explanatory power of the eight ratio variables is

$$\begin{aligned} R_{\mathrm{s}} = {} & \mathrm{a}_0 + \mathrm{a}_1\mathrm{EP}_{\mathrm{s}} + \mathrm{a}_2\mathrm{BP}_{\mathrm{s}} + \mathrm{a}_3\mathrm{CP}_{\mathrm{s}} + \mathrm{a}_4\mathrm{SP}_{\mathrm{s}} + \mathrm{a}_5\mathrm{REP}_{\mathrm{s}} \\ & + \mathrm{a}_6\mathrm{RBP}_{\mathrm{s}} + \mathrm{a}_7\mathrm{RCP}_{\mathrm{s}} + \mathrm{a}_8\mathrm{RSP}_{\mathrm{s}} + \varepsilon_{\mathrm{s}} \end{aligned} \tag{8.8}$$

In the context of a rolling quarterly backtest of the potential benefit of using this type of ratio-based stock return forecast to improve portfolio performance using a mean—variance optimizer, BGMTX creates a time series of rolling one-quarter-ahead return forecasts from the estimated regression coefficients from Eq. (8.8).

[20] *Security analysis*, Graham and Dodd (1934), is generally credited with establishing the idea of *value investing*. Graham and Dodd influenced Williams (1938), who made particular reference to their low P/E and net current approaches in *The Theory of Investment Value*. In turn, Williams (1938) influenced Markowitz's thoughts on return and risk as noted in Markowitz (1991). Over the past 25 years, value-focused fundamental analysts and portfolio managers have expanded their value measures from primarily price relative to earnings and price relative to book value to include also price relative to cash flow and even price relative to sales. The choice of the eight fundamental variables in BGMTX reflects this expansion in focus, especially the expansion to include cash and sales ratios.

They use four quarters of past coefficients estimates (four sets of quarterly estimates) as the basis for a relative weighting of the eight value ratios. For each quarter-ahead return forecast, BGMTX develops a relative weighting by first modifying the coefficient estimates as described below to reflect significance and extreme values, then averaging the modified coefficients from the past four quarters, and finally normalizing the averaged coefficient values.

Stone and Guerard (2010) replicate the BGMTX forecast procedure: (1) to test performance after the publication of the model in 1993, (2) to expand the time period and sample size for the model performance potential evaluated, and (3) to resolve questions of whether the apparent performance is at least in part a return for risk or possibly a de facto yield tilt or possibly even from other return impact variables. The question of a de facto risk tilt is especially pertinent because the Fama–French return model includes BP as one of its three risk variables and BP is also one of the BGMTX return forecast variables.

The illustration of the matched control methodology to isolate well forecast performance from risk and other distortions is based on Stone and Guerard (2010). As a post publication test of the original model, the only change that Stone and Guerard (2010) made to the BGMTX return forecast procedure itself is to forecast monthly returns in a rolling month-to-month framework rather than forecasting quarterly returns in a rolling quarter-to-quarter framework.

Detailed below is the step-by-step forecast procedure summary as adapted in Stone and Guerard (2010) for a rolling month-to-month forecast.

1. *Regression coefficient estimation*. With a two-month delay, estimate each month for ten months back the regression coefficients $\{a_0, a_1, \ldots, a_8\}$ of Eq. (8.8) above.
2. *Coefficient modification*. Adjust/modify regression coefficients a_1 to a_8 in each month to reflect significance and/or extreme values in two ways:
   (a) Any coefficient with a t-value ≤ 1.96 is set equal to zero.[21]
   (b) Extreme positive values are truncated.
3. *Normalized average*. Average the last ten months adjusted coefficient values and normalize these averages to determine relative weights that sum to one. Let w_i denote the normalized forecast coefficient for the ith value variable, $i = 1, \ldots, 8$. The $\{w_i\}$ sum to one.
4. *Update ratio variables*. For each stock in the sample, update the eight value ratios using the current stock price and financial statement variables as reported in Compustat from the "most recent" (at least 2-month back) annual financial statement and current stock prices.
5. *Compute forecasted return*. Use the normalized weights from step 3 and the updated ratios from step 4 to obtain a month-ahead return forecast. If FR_s denotes

[21]When regression coefficients with t-values ≤ 1.96 are made equal to zero, there are no negative coefficients regardless of significance.

the forecasted return for stock *s*, then the formula for the forecasted return for stock *s* is the weighted average of the eight value ratios, i.e.,

$$\begin{aligned} FR_s = {} & w_1EP_s + w_2BP_s + w_3CP_s + w_4SP_s + w_5REP_s + w_6RBP_s \\ & + w_7RCP_s + w_8RSP_s \end{aligned} \tag{8.9}$$

The forecast formula in Eq. (8.9) is similar to the cross-sectional return regression except that:

1. The regression error term is dropped.
2. There is no intercept coefficient.
3. The regression coefficients in Eq. (8.8) are replaced by the 10-month average of significance-adjusted, outlier-modified, and normalized past coefficient estimates for each variable.

8.7 Control Variables

8.7.1 Control Constraints

To assess the performance potential return forecast, it is essential to eliminate any impact from systematic risk, tax effects, or other nonmodel variables such as growth and profitability that could conceivably be the source of apparent performance value. Exhibit 8.1 lists a set of risk, tax, growth, and profitability variables that are candidate control variables.

8.7.2 Risk Controls: β, BP, and Size

The first three variables listed in Exhibit 8.1 are the three Fama–French risk variables: beta, book–price, and size. The ex ante beta value used in this study was based on a rolling update using three past years of monthly risky returns (return in excess of the monthly T-bill rate in that month) relative to the risky return on the CRSP index.

The *book–price ratio BP* is the book value per share divided by price per share. The ex ante book–price value is computed using the book value from the most recent financial statement lagged at least two months to allow for the financial statement data to be public information. The price per share is the last closing price in the prior month. Using BP as a risk variable is consistent with the Fama–French risk modeling but conceptually different from the Graham–Dodd use of BP as one of the value ratios that can indicate relative misvaluation of otherwise comparable companies. Given that BP is one of the eight ratio variables in the BGMTX forecast model, the critical performance question is whether the contribution of BP is a risk effect in disguise or whether it is an indicator of value potential beyond any systematic risk. Rather than the either–or extremes of being either all risk ala Fama–French or all performance value ala fundamental value-focused analysts, the reality is almost certainly a combination of risk and value potential with the critical performance question being the relative amount of risk and value beyond risk at a given point in time. The relative amount of each effect in a cross section of performance-ranked return predictions is almost certain to vary across time. For a researcher trying to assess true value performance potential, resolving these relative contributions is a difficult problem. As we discussed further in Sect. 8.8 illustrating the imposition of risk isolating control variables, it is a very difficult problem to resolve via the conventional multivariate regression assessment but more treatable by the matched control methodology.

The size variable *S* is simply the market value of outstanding equity, the price per share times the number of outstanding shares. The ex ante value used in this study is based on the price per share at the close of trading in the prior month. While the measurement of size has the least measurement error of the three risk variables, the cross-sectional size distribution is perverse in the sense of having a large number of relatively small cap companies and a small number of very large cap companies. To produce a less extreme size distribution that mitigates the extremely large weight given to the small number of very big companies, Fama–French and other researchers have used *the log of company size* as the size measure in assessing the ability of size to explain the cross section of stock returns. This rather arbitrary nonlinear transformation mitigates but does not cure the heteroscedasticity problem. As in a cross-sectional regression, it matters how size is measured when imposing an equal-size constraint using the power optimizing transformation detailed in Sect. 8.5. In particular, imposing a size control that makes every portfolio in the cross section in a given month have the same average size can mean reassigning a very large company to several portfolios in order to satisfy the equal average size constraint. As in a cross-sectional regression, using the log of size mitigates but does not really cure this size distortion. An alternative used in this study was the creation of a relative size variable. "Relative size" is obtained in a given month by dividing all companies by the size of the largest company. Thus, the relative size variable puts all companies on the interval (0,1). The range in cross-sectional variance is comparable to the range and cross-sectional variance for beta, financial leverage, and growth and clearly less than the range and cross-sectional variance in other control variables such as BP, EP, and DP.

8.7.3 Tax Controls: DP, EP, and FL

The ex ante dividend yield variable DP is an annualized value of the most recent quarterly dividend per share at least 2-months back divided by the share price at the end of the prior month. Because dividends change slowly, there is very little measurement error in using the ex ante dividend as a predictor of the future dividend. As with BP, most of the uncertainty in DP arises from changes in the price per share. For this reason, the cross-sectional correlation between BP and DP is high.

The primary tax control is DP. With the dividend yield control, every portfolio in the cross section will have the same portfolio average dividend yield. Hence, the DP control means that every portfolio has the same ex ante expectation for ordinary income. With the same dividend yield, the cross section of realized returns becomes a cross section of realized capital gains. Any variation in the dividend–gain mix over the cross section is a capital gain effect. If beta were also controlled, the ex ante CAPM expectation is a flat cross section.

Given a normalized set of weights for the BGMTX return forecast, the higher predicted returns tend to correspond to stocks with higher values of the four current value ratios, BP, EP, CP, and SP. These are all correlated with dividend yield and thus the concern that the apparent return performance may actually be a dividend yield tilt in disguise. Thus, apparent before-tax performance would be significantly reduced or eliminated if returns were put on an after-tax basis.

Adding a control for the earnings yield EP to the DP control tends to improve the ability of the ex ante DP variable to be a good control for the dividend-gain mix. When each portfolio in the cross section has the same average value of both EP and DP, each portfolio has the same ex ante dividend payout ratio, i.e., the same portfolio average value of the dividends–earnings ratio. To the extent that the dividend payout ratio characterizes dividend policy, the combination of the EP and DP controls together means that each portfolio in the cross section has the same portfolio average dividend payout policy.

One further comment on the effect of the DP control alone and especially in combination with the EP and FL controls pertains to the interaction with both size and beta. Stocks having a high dividend yield and high earnings yield tend to be larger companies with lower than average beta values. Hence, imposing the DP control alone and especially the DP and EP controls together tends to move larger and lower beta stocks to lower-ranked portfolios and vice a versa, to move smaller and higher beta stocks to higher-ranked portfolios.

It is common to talk about a value/growth trade-off with the assumption being that high value tends to mean lower growth and vice a versa. **Given validity to the value/growth assumption, value controls like BP, EP, and DP are also de facto growth controls**. Exhibit 8.1 lists two growth control variables: 5-year past sales growth and sustainable growth. When used in addition to the risk and tax controls, it is reasonable to assume that these two controls are simply refining the already established growth control associated with the risk and tax controls. This point is discussed further after illustrating the use of the risk and tax controls.

The financial leverage control FL is the ex ante percentage of nonequity financing. It is measured as one minus the book equity per dollar of total investment, both values being from the most recent annual financial statement at least 2-months back. Aside from preferred stock, FL measures the percentage of total investment provided by debt financing.

Financial leverage has been included with DP and EP as a tax control. Rather than controlling for the dividend–gain mix, FL is designed to reflect the corporate tax shield associated with debt financing and thus a corporate tax performance impact associated with more return to shareholders and less to the government. Like many controls, FL reflects more than just the tax shield of debt financing. Use of debt involves an increase in both refinancing and interest rate risk. One source of adverse changes in interest rates is a change in inflationary expectations. Hence, the stock-specific FL control has an element of both company and macro risk control in addition to reflecting any valuation effect of the corporate tax shield.

Another potentially important role for the FL control pertains to industry exposure. There is considerable variation across industries in the relative use of debt financing. Thus, controlling for financial leverage tends to be a de facto control on the variation in industry mix across portfolios. A check on industry membership over the uncontrolled cross section of forecast-ranked portfolios compared to cross sections with the FL control imposed indicates a clear but less-than-perfect tendency for the FL control to reduce well concentrations of some industries in subsegments of the uncontrolled cross section.

8.8 Using Control Variables to Isolate Performance Potential

Section 8.8 uses the control variables defined in Sect. 8.7 and the return forecast model summarized in Sect. 8.6 to provide a concrete illustration of the power/efficiency benefits of the matched control methodology for a full sample assessment of the performance potential of a stock return forecast. The purpose here is not to establish value for the BGMTX forecast model per se but rather to use an actual return forecast to illustrate the power/efficiency benefits of the matched control methodology.

This illustration emphasizes the major design decisions that impact statistical power and efficiency.

8.8.1 Alternatives to the Full Sample, Relative Rank-Ordering Framework

From the viewpoint of having a high-power statistical assessment of the performance potential of a stock return forecast, the most important decision is the selection of the assessment framework. One alternative to the full sample relative rank ordering advocated here is the use of the forecast to select a stock portfolio that is then compared to a reference index benchmark. The BGMTX forecast evaluation was based on the ability of the return forecast along with a risk forecast to generate a frontier of mean–variance optimized portfolios that outperformed a reference index. One problem with this approach is ambiguity with respect to the relative performance contribution of the return forecast versus the risk forecast although good sensitivity analysis can reduce this ambiguity.[22] The significant power disadvantage is limiting the forecast evaluation to the selected stocks, i.e., the stock subsample that has the best expected risk-adjusted return. Not assessing performance potential for the full sample means a loss of information and thus a loss of statistical power.

For the management of active mutual funds and for hedge funds, typical use of return and risk forecasts is not to generate a mean–variance efficient frontier but rather to use a stock return forecast as input to an index–tilt portfolio selection model that seeks to maximize the increase in expected return relative to a benchmark index subject to constraints on tracking error and a maximum tilt away from tracking error-related style variables such as beta, size, value/growth, industry, and country.[23] For organizations operating in an index tilt environment, the standard backtest performance potential assessment is to generate a time series of forecasts over a pertinent past time period and then evaluate the average performance improvement relative to the benchmark or possibly relative to another forecast or even just the past performance of the fund. The comparative assessment of alternative forecast selection approaches is often termed a "performance derby."

While assessing return forecast benefits in the context of the portfolio tilt environment in which the forecast is to be used is clearly an essential step in evaluating the performance potential of a return forecast, such a backtest assessment of constrained portfolio selection is a logical follow-on after first establishing how well the forecast performs in a large sample backtest, at a minimum how well the forecast performs in terms of ability to identify misvalued stocks across at least all the stocks in the benchmark index plus any stocks that are candidates for replacement of benchmark stocks.

There are two problems with skipping a full-sample, relative rank-ordering performance assessment and only assessing benchmark tilt performance. As in the case of evaluating a return forecast via mean–variance portfolio selection, the tilt to a relatively small subset of the stocks in a benchmark index means loss of potential sample information and thus loss of power. The typical benchmark tilt is almost always a small departure from the benchmark, for instance, a 20 % tilt is generally viewed as relatively large with significant tracking risk. Thus, for an S&P 500

[22]Michaud (1989, 1998) recognizes that uncertainty about both the return and risk forecasts and other portfolio selection parameters is a source of risk/uncertainty in addition to the inherent uncertainty risk of investment and has formalized a very sophisticated resampling simulation to structure very thorough sensitivity analysis.

[23]See Grinold and Kahn (2000) for a thorough description of the MCI-Barra active tilt frameworks. See Menchero, Korozov, and Shepard (2010) for an updated version of the MCI-Barra equity risk modeling that includes both industry and country factors, a global equity risk factor, and additional style factors for value, size, momentum, etc.

benchmark, the performance of the portfolio is typically more than 80 % benchmark and at most a 20 % tilt.[24] The effective comparison sample is about 100 stocks either predicted to have the best expected return relative to tracking error for overweighted stocks or the worst expected return relative to tracking error for underweighted or excluded stocks. **By focusing on a small subset of the pertinent stocks that is primarily the subset that is predicted to be the extreme best and worst stocks, a benchmark tilt comparison excludes information on the ability to rank order the rest of the pertinent stock universe**. Compared to a full sample relative rank-ordering assessment, just using a benchmark tilt assessment is a low power relatively uninformative performance potential assessment.

The assertion of being uninformative pertains especially to the second problem with using a benchmark tilt assessment to evaluate a stock return forecast, namely, mixing any forecast performance value with the effect of predictions of tracking error and of style alignment/misalignment. Mixing the effect of a return forecast with tracking error predictions and style and industry alignment/misalignment obfuscates information about the forecast itself. While extensive statistical and sensitivity analysis can help separate forecast performance from other factors,[25] the clear best solution to having a high-power assessment of forecast performance potential is to isolate completely return performance from all other return impact variables. The primary function of the matched control embellishment of the relative rank ordering is to ensure a well-isolated return forecast. Use of the power optimizing reassignment programs like that formulated in Sect. 8.5 ensures that power is optimized.

Within the full sample, relative rank-ordering framework, there are alternatives to the power optimizing matched control methodology advocated here. The most common is to use multivariate regression to explain realized returns by a combination of the return forecast and other known return impact variables such as the control candidates developed in Sect. 8.7. Another alternative is to use an endogenous APT to remove all statistically identifiable systematic variation from the return as illustrated in Guerard, Gültekin, and Stone (1997). The merits of matched controls compared to these two alternatives are addressed later in the context of illustrating and evaluating the matched control methodology.

[24]The magnitude of a tilt is defined as the absolute value of the difference in the relative weighting of stocks in the benchmark and the tilt portfolio. For instance, if a stock with a weight of 0.5 % is increased to 0.7 %, the tilt change is 0.2 %. If a stock with a relative weight of 0.15 % is excluded completely, the tilt change is 0.15 % to the tilt. The overall tilt percentage is the sum of all the tilt change percentages.

[25]The resampling simulation approach set forth in Michaud (1989, 1998) is again pertinent here.

8.8.2 Stepwise Imposition of Control Constraints: Procedure Overview

In a control-based assessment of how apparent forecast performance value is distorted by interaction with control variables, the starting point is a collection of rank-ordered return forecasts with no controls imposed. Inputting the no-control cross section to the mathematical assignment program for a given set of control variables produces an output cross section in which each portfolio is now matched on the specified set of controls. Comparing before and after cross sections and noting any changes in the cross section enable a forecaster/researcher to assess to what extent apparent performance potential has been distorted by one or more of the control variables. Or, in the case of no significant change, a forecaster/researcher knows that the given set of controls is not distorting apparent performance potential.

Adding controls in a stepwise fashion enables a researcher to explore how the initial rank-ordered cross section changes by systematically removing the effect of a control variable or combination of control variables. This stepwise exploration of how the return dependency changes with changes in combinations of control variables is generally very informative. Because the primary concern here is correcting apparent performance for distortion from risk and tax effects, the stepwise assessment of control impacts focus primarily on cross sections for six sets of controls summarized below:

1. No controls: the initial rank ordering
2. Individual risk controls: beta, book-to-market, and size as individual controls
3. Three risk controls together: beta, book-to-market, and size together
4. Three tax controls together: the earnings–price ratio, the dividend–price ratio, and financial leverage together as a combination control for the dividend–gain mix and other tax effects
5. The combination of risk and tax controls: the three risk and the three tax controls together, six control variables in all.

After in-depth assessment of the effect of risk and tax controls, the impact of growth and profitability controls is assessed. Finally, by removing the effect of the four value ratios BP, EP, CP, and SP, we assess the relative contribution of the four value ratios and the four relative value ratios to forecast performance.

8.8.3 Study Sample and Time Frame

The backtest study period is January 1967 through December 2004. Developing a return forecast for every stock in the backtest sample for January 1967 through December 2004 produces a time series of 456 monthly return forecast cross sections.

The data sample is all nonfinancial common stocks in the intersection of CRSP and Compustat with a book value in excess of $20 million that are included in CRSP for at least three years with monthly return data necessary to compute a 3-year rolling beta and in Compustat for at least five years with all necessary financial statement data. The table below summarizes by year the number of companies in the 1967–2004 backtest study sample.

Year	#Stocks	Year	#Stocks
1967	198	1986	1660
1968	324	1987	1632
1969	422	1988	1580
1970	564	1989	1621
1971	901	1990	1644
1972	966	1991	1671
1973	1058	1992	1742
1974	1108	1993	1845
1975	1037	1994	1921
1976	1329	1995	2003
1977	1495	1996	2057
1978	1651	1997	2193
1979	1701	1998	2238
1980	1703	1999	2331
1981	1757	2000	2284
1982	1734	2001	2256
1983	1698	2002	2305
1984	1714	2003	2318
1985	1676	2004	2238

Because of the sparseness of the Compustat database in the 1964–1966 5-year start-up period required for control variables such as 5-year sales growth, there are only 198 companies in January 1967 and only 324 companies in January 1968. The table shows that the forecast sample size grows rapidly. From 1971 on, there are more than 900 companies in the forecast sample growing to more than 2000 companies by 1995.

The fact that the sample size shows little growth from the 2003 stocks in January 1995 to the 2238 stocks in January 2004 indicates that the large number of new IPOs after the mid-1990s is not producing an increase in the number of sample companies. The fact that our sample does not exhibit the same growth as the cross time increase in publicly listed companies shows that the combination of requiring 5 years of past financial statement data plus the minimum book value restrictions means that we are studying primarily larger more mature companies.

8.8.4 Key Efficiency/Power Design Decision: The Number of Fractile Portfolios

In panel studies using rank-ordered cross sections, many return dependency assessments rank stocks into deciles and in some studies only quintiles. However, in each month of 1967–2004, stocks were ranked into 30 fractile portfolios. As already discussed, having more fractile portfolios pertains to the power/efficiency trade-off. Most of the efficiency benefits of measurement error and omitted variable diversification are accomplished with 20 or fewer stocks in a fractile.

It is pertinent to expand on the greater power benefits of having more fractile portfolios. First, grouping stocks into a fractile portfolio and using the portfolio-weighted average to represent the portfolio value of variables tend to average away information while averaging away measurement error, especially in the tails of the distribution. Second, cross-sectional plots like those in Exhibits 8.2, 8.3, and 8.4 of key performance indicators such as average realized return and realized standard deviation are more useful when there is a high density of data points. Third, when assessing the cross-sectional dependence of realized returns and realized standard deviation cross sections on the return forecast, both efficiency and power are increased from more observations. In particular, regression estimation and related hypothesis testing have much greater statistical power when there are at least 20 observations. Fourth, adjacent portfolios in a control-matched return cross section can be merged together and preserve the control matching without having to resolve the control matching optimization program. For instance, combining adjacent three tuples in the 30-portfolio cross section produces a cross section of matched deciles as done in Sect. 8.8.9. Thus, **it is methodologically better to error on the side of too many portfolios in a cross section rather than too few**.

8.8.5 The Impact of Individual Risk Controls

Exhibits 8.6 shows cross-sectional plots of average realized return and realized standard deviation versus portfolio rank for no controls and for just a beta control for risk. Exhibit 8.3 presents return and risk cross sections for just a size control and just a BP control.

Plot 2.1 is a cross section showing average realized return for the rank-ordered cross section with no controls. If there were no forecast information in the return prediction, the plot would be a random scatter about the overall average return. In contrast to a random scatter, Plot 2.1 shows an overall tendency for realized return to increase with an increase in predicted return.

The rate of increase is clearly nonlinear. For portfolios 1–10, the cross section is noisy and relatively flat. For portfolios 10–20, the realized return increases at a steady rate. For portfolios 20–30, the average realized return tends to increase at an accelerating rate with the largest increases being for portfolios 28, 29, and 30.

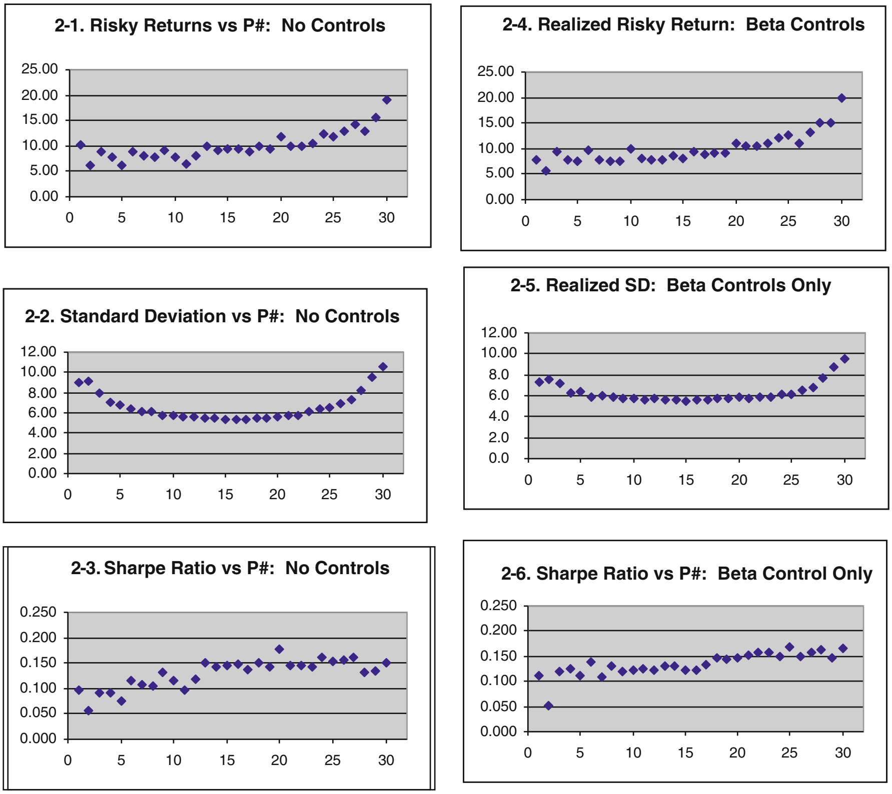


Exhibit 8.2 Risky returns, SD, and Sharpe ratio vs. P# (portfolio #): no controls compared to only a beta control

The realized return range is large and economically significant. The annualized realized return difference between portfolios 30 and 1 is 8.8 %. The difference between the realized return for the upper quintile (average for the top six portfolios) and the realized return on the bottom quintile (average return for the lowest six portfolios) is 6.5 %.

In effect, before imposing controls, Plot 2.1 indicates that the return forecast has limited ability to rank order return performance for the bottom third of the sample other than identifying the bottom third as inferior to the rest of the sample. For portfolios 10–30, the return forecast tends to rank order on average portfolio-level realized risky return. The relative rank-ordering ability is especially good for the top third of the sample, portfolios 21–30. The apparent ability to predict realized risky return improves with portfolio number and seems especially good for the top two portfolios.

The vertical axis in Plot 2.1 is average realized risky return. There is no correction for possible variation in risk. Before imposing controls for risk or otherwise correcting for any portfolio-to-portfolio variation in risk, basic insight on risk variation is provided by measuring the cross-time variation in realized risk. Plot 2.2 shows the cross-sectional dependence of realized standard deviation on portfolio number for the case of no controls. For each portfolio number, the realized standard deviation is computed in accord with the definition by taking the square root of the mean squared deviation of each annualized monthly return from the long-run average return. The standard deviation cross section in Plot 2.2 is not a random scatter. It is also not the steady increase implied by the assumption that higher predicted returns arises from selecting progressively higher risk stocks. The very steady portfolio-to-portfolio variation pattern is not only nonlinear but also nonmonotonic. If an increase in predicted return were associated with a systematic increase in realized standard deviation risk from either systematic or unsystematic sources, the cross section of realized standard deviations should be increasing with an increase in predicted return. Any interaction between risk and predicted return is more complex than a simple linear association.

The SD cross section in Plot 2.2 is much smoother than the return cross section return in Plot 2.1. Compared to the 8.8 % return range, the SD cross section has a smaller range of just 5.19 %.[26] Most of this range is from the relatively high realized standard deviations for the very low return forecasts and the very high-return forecasts. For the three inner quintiles, portfolio 7 through portfolio 24, all realized standard deviations are within a range of just 0.70 %.

The very smooth, nonmonotonic SD cross section raises questions. One implication is that the very high and the very low return forecasts have greater realized SD risk. One question is whether the greater realized standard deviation risk arises from greater systematic risk or from unsystematic uncertainty or even possibly greater forecast uncertainty for the extreme forecasts. Another related question is whether the higher returns in the upper quintile justify the greater SD. From the viewpoint of both mean–variance efficiency and tracking error control, whether the source is systematic, unsystematic, or greater forecast error is a very important information. Imposing risk control first for beta, size, and BP individually and then in combination can help answer these questions.

8.8.6 CAPM Performance Assessments

Plot 2.4 is a cross section showing average realized return for the rank-ordered cross section with a control for just beta risk but with no other controls imposed. Imposing the same beta control means that every portfolio in the beta-controlled cross section in Plot 2.4 has the same population average value of beta. Since the

[26]See Appendix 8.1 for the data to compute the ranges of realized risky return and standard deviations. Appendices 8.2 to 8.6 provide pertinent data for each of the 30 fractile portfolios for the cross sections with other control variables.

rolling three-year risky return beta was calculated relative to the sample average, the ex ante sample average beta value is one in every month.[27]

The textbook *Treynor Performance Measure* is realized risky return divided by the beta of the portfolio. Thus, with all portfolios having a beta of one, **the average realized return cross section in Plot 2.4 is also a Treynor performance cross section**.

Compared to the no-control cross sections, the beta controls tend to smooth somewhat the return cross section but with very little systematic change. In contrast, the beta controls smooth the SD cross section and significantly reduce the range. The overall range is reduced from 5.19 % with no controls to 4.01 %. More significantly, the SD range for the 18 interquartile portfolios is reduced from 0.70 % for no controls to just 0.55 % with beta controls. Given the smoothing and especially the reduced range, the beta control seems to do a good job of correcting portfolios in the three inner quartiles for variation in realized standard deviation risk.

The fact that beta controls reduce but do not eliminate the greater standard deviations for both the very high and very low forecasts is evidence that some of the SD increase for these portfolios is beta related. However, for the highest predicted returns, there seems to be more to realized SD uncertainty than just beta.

The widely used *CAPM alpha* is the realized risky return less beta times the average risky return on the market index portfolio. When beta is one for every portfolio, we have

$$\boldsymbol{\alpha_p = R_p - \beta_p\,(\text{risky index return}) \rightarrow \alpha_p = R_p - (\text{risky index return})\,,\ p = 1, \ldots,\ P.}$$

When beta is the same for every portfolio in the cross section in Plot 2.4, the CAPM alpha is just an additive constant subtracted from the realized risky returns of each portfolio in the cross section. Thus, **to within an additive constant for the average realized risky return on the market index portfolio, the average realized return cross section in Plot 2.4 with all portfolios having a beta of one is also a CAPM alpha performance cross section**.

The third standard performance measure is the Sharpe ratio. Like the Treynor ratio, the Sharpe ratio assesses risky return relative to the associated risk. Rather than risky return per unit of beta risk, the Sharpe ratio is risky return per unit of standard deviation risk. Plots 2.3 and 2.6 show the cross section of realized Sharpe ratios. As expected given the wide range of realized risky returns and relatively smaller range of realized standard deviations, the overall trend for the Sharpe ratio without any controls and especially with the beta control is a tendency to increase with an increase in predicted return. Given the smoothing from the beta control and especially given the reduced range for realized standard deviations, the beta-controlled Sharpe ratios are smoother and more nearly monotonically increasing than the uncontrolled cross section.

---

[27] Relative to the CRSP index, stocks in the backtest sample had a lower beta, generally about 10 % lower. While the backtest sample excluded generally low beta financial stocks, it also tilted toward larger, more mature companies because of the requirement of inclusion in both Compustat and CRSP for at least five years. This maturity tilt is the reason for the somewhat lower beta values for the backtest sample than for the overall CRSP sample.

Compared to the Treynor and the CAPM alpha, the relatively large realized standard deviations for the highest predicted returns mean that the beta-controlled Sharpe ratio performance assessment is less favorable for the highest predicted returns than either the Treynor or the CAPM alpha.

8.8.7 *The Impact of Size and BP Risk Controls*

Afterconsidering the CAPM beta-controlled performance, the next logical step is to evaluate the impact of the two remaining Fama–French risk variables—size and BP. Assessing the risk impact of BP is especially pertinent since it is one of the eight predictor variables in the BGMTX return forecast.

Plots 3.1–3.3 in Exhibit 8.3 show the cross sections of realized return, realized standard deviations, and realized Sharpe ratios when a size control is imposed.

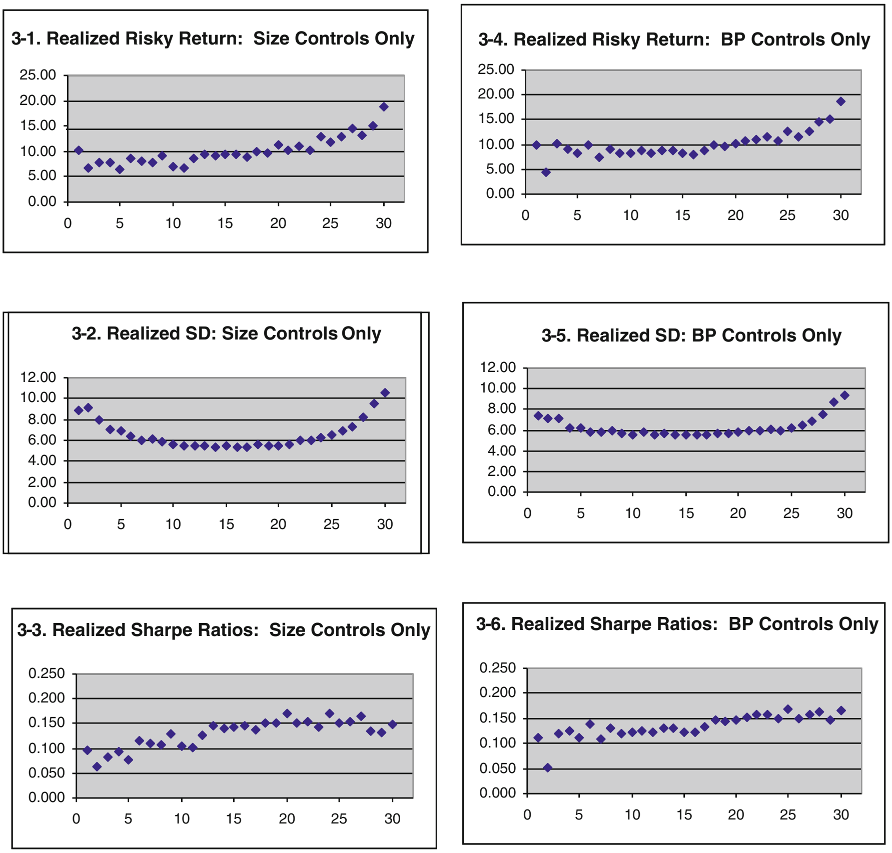


Exhibit 8.3 Risky return, SD, and Sharpe ratio vs. P# (portfolio #) for a size control only and a BP control only

Contrary to what one expects from the tendency of size to be negatively correlated with the four value ratios BP, EP, CP, and SP, the imposition of the size control alone has very little impact relative to the uncontrolled cross sections other than a slight smoothing of each cross section and a modest reduction in the realized return for the highest return forecasts.

Plots 3.4–3.6 in Exhibit 8.3 show the cross sections with just a control for BP. Despite the fact that BP is one of the eight variables in the return forecast model, eliminating any impact of BP on the cross sections has a relatively modest impact on the range of realized returns.

Imposing just the BP control tends to smooth the three cross sections, especially for portfolios ten and higher. Consistent with being a risk variable, imposing just the BP control reduces the range of realized standard deviations.

8.8.8 Imposition of Combinations of Risk and Tax Controls

Exhibits 8.4 and 8.5 repeat plots of the return cross section and SD cross section for no controls and then show the return and SD cross sections, respectively, for three key combinations of controls:

1. The Fama–French risk controls: beta, size, and BP
2. Three tax controls: DP, EP, and FL
3. The combination of the three Fama–French risk controls and the three tax controls

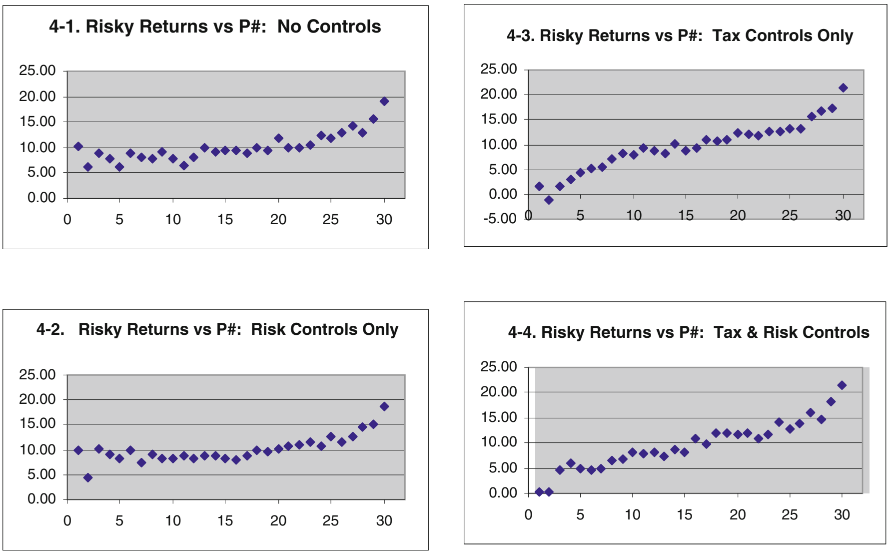


Exhibit 8.4 Risky returns vs. portfolio number: risk only, tax only, and both risk and tax controls together

Plot 4.2 of Exhibit 8.4 summarizes the realized return cross section when all three Fama–French risk controls, beta, size, and BP, are imposed together. Recall that making these three variables into control variables means that the transformed cross section has shifted securities so that each portfolio in the cross section has the same portfolio average value of each of these three variables. There is no portfolio-to-portfolio variation in the value of beta, size, or the book–price ratio. Thus, in each month, these three variables will have the same contribution to portfolio return. Realized risky return is now well-isolated from any differential impact from any of these three risk variables. The portfolio-to-portfolio variation in return must arise from the forecast variables other than the now-controlled book–price ratio, possibly taxes, or other return impact variables but not from beta, size, or BP.

Comparing Plots 4.1 and 4.2 shows a similar range and pattern but with much less portfolio-to-portfolio variation in average realized return. The net effect of the three risk controls is to smooth the curve without changing the overall nonlinear pattern or the range of realized risky returns. Moreover, the smoothing effect makes the nonlinearity much more pronounced. For the no-control plot and especially risk control plot, the cross section of average realized returns is flat to slightly declining for portfolios 1–15. For portfolio 15 on, the cross section has a steady monotonic increase with the rate of increase being the greatest for the top three portfolios.

Plot 4.3 of Exhibit 8.4 summarizes the realized return cross section with three tax controls: EP, DP, and FL. Making each portfolio in the cross section have the same average value of the dividend price ratio by itself tends to ensure that the percentage of return realized as dividends is the same as the percentage of return realized as capital gains. When both EP and DP are the same in every portfolio, this amounts to each portfolio having the same dividend payout ratio, which is an additional control on dividend policy. Financial leverage has been included as a tax control to reflect the tax deductibility of corporate interest payments. The combination of having the same earnings price ratio and therefore almost the same average earnings for each of the portfolios plus the same percentage of debt means roughly the same average percentage of earnings are shielded from taxes. Financial leverage also tends to reflect both company debt capacity and exposure to interest rate risk and may reflect some performance and risk beyond the three control variables that we have characterized as "risk controls."

Comparing Plot 4.3 with Plots 4.1 and 4.2 indicates significant changes in the cross section of realized returns compared to no controls and especially compared to the cross section with all three risk controls together. The largest changes are for portfolios 1–15. Realized returns are reduced on average and rank-ordered much better for portfolios 1–15 than for the plot for no controls or also for all three risk controls together. The overall monotonic increase is now much steadier and more nearly linear. **Tax effects clearly exhibit very significant systematic variation over the cross section of forecast rank-ordered portfolios compared to both the uncontrolled and the risk-controlled cross sections!**

Plot 4.4with both risk and tax controls is similar to Plot 4.3. Adding risk controls to the tax controls does not significantly change the cross-sectional return plot. Having controls on both DP and β together for the combination of risk and tax

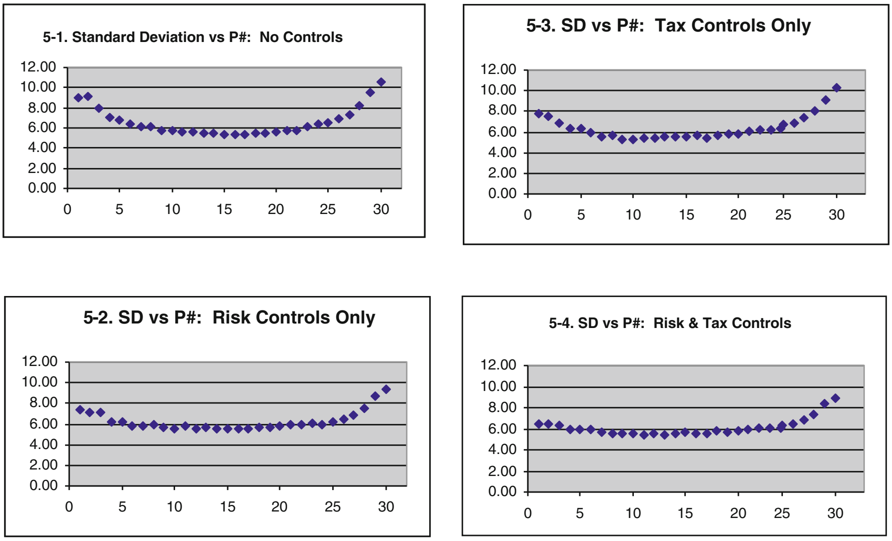


Exhibit 8.5 Standard deviation vs. portfolio number: risk only, tax only, and both risk and tax controls together

controls tends to ensure even more completely that the percentage of returns realized as dividends and as capital gains is the same over the cross section.[28]

Exhibit 8.5 contains plots of realized standard deviation versus portfolio number for no controls, for all three risk controls together, for just the three tax controls together without any risk controls, and then for all three risk and tax controls together. All four plots are nonlinear and nonmonotonic. Plot 5.1 with no controls has a range of realized standard deviation values from a low of about 5.5 % in the

[28]It is easy to check how well DP alone and DP and β in combination actually control for cross-sectional variation in the dividend–gain mix. For the forecast rank ordering in this study, in all time periods of 5 years or longer after 1972, the DP control alone does a good job of controlling for cross-sectional variation in the ex post dividend–gain mix. The term "good job" means that the average ex post dividend–gain ratio in each portfolio is very close to the sample average with no systematic variation over the cross section. Controlling for DP and EP together improves the control for variation in the dividend–gain mix by eliminating portfolio-to-portfolio variation and making most portfolios very close to the average. Controlling for DP and β in combination improves the control since high beta tends to be lower dividend payout. Likewise controlling for DP and size in combination improves the control for the dividend–gain mix for similar reasons, namely, the fact that small size tends to be higher beta and often zero or token dividend payout. Thus, the combination of tax controls and risk controls together improves on the tax controls alone in terms of ensuring very little portfolio-to-portfolio variation in the dividend–gain mix, especially for all time periods after 1972. The main caveat is slighter greater variation about the average dividend–gain mix for the three lowest-ranked and the three highest-ranked portfolios. This greater variation about the mean for the low-ranked and high-ranked portfolios is consistent with the much greater realized standard deviation for these portfolios as shown in Exhibit 8.5 as well as the much greater positive skewness for the three highest-ranked portfolios.

middle of the cross section to a high of 11.5 % for portfolio 30, a high-low range of 6 %. The portfolio-to-portfolio changes are remarkably smooth. After portfolio 2, realized standard deviations first smoothly decrease toward the middle of the cross section and then smoothly increase at an accelerating rate with the highest realized standard deviation occurring for portfolio 30.

As expected, adding the three risk controls in Plot 5.2 of Exhibit 8.5 tends to reduce realized risk variation as reflected in the reduced range of realized standard deviations from 6 % to less than 3.5 %. Moreover, realized standard deviation varies by no more than 1 % from portfolio 4 to portfolio 20. Most of the increase in realized standard deviation at the low and high end of the cross section is attributable to skewness, negative skewness for the low end, and significant positive skewness for the high end.

Plot 5.3 for tax controls only is similar to Plot 5.2 for just risk controls except for slightly higher realized standard deviations at the low and high extremes and a slightly greater asymmetry. The fact that both risk and tax controls have similar effects in terms of controlling for realized standard deviation risk is surprising. It suggests risk control impacts from some combination of the dividend–gain mix and possibly financial leverage risk. Plot 5.4 with both risk and tax controls together supports this conjecture of risk control benefit from the three tax controls beyond the risk control provided by beta, book–price, and size. For portfolio 1–21, the cross section of realized standard deviations varies by just a little more than 1 %. For these 21 portfolios, the combination of the conventional risk variables plus the tax controls does an excellent job of controlling for realized risk as measured by realized standard deviation.

As indicated by the skewness data in the Table in Appendix 8.6, the increase in realized standard deviation in Plot 5.4 is well explained by the corresponding increase in significantly positive realized skewness for the highest return portfolios.[29]

Exhibit 8.6 shows cross-sectional plots of the Sharpe ratio for the four control sets. As a synthesis of the respective return and standard deviation plots, they show that the overall increase in returns is outweighed by the very modest increase in realized standard deviations. Hence, the plots in Exhibit 8.6 indicate significant performance potential for the basic BGMTX return forecast framework, with the most pertinent Sharpe ratio cross section being Plot 6.4 since it eliminates distortion from both risk and tax effects.

[29] A plot of realized semi-standard deviation for the case of all risk and all tax controls together is flat for the top 25 portfolios in the cross section, strong support for the assertion that the greater standard deviation for the top four portfolios is primarily a positive skewness effect and not downside uncertainty.

The primary purpose of this control matching example is to illustrate the benefits of using matched controls in assessing forecast performance potential rather than to establish value to the eight-variable BGMTX forecast model.[30]

8.8.9 *Stepwise Imposition of Risk and Tax Controls: High-Minus-Low Differences*

Exhibit 8.7 summarizes high-minus-low returns for major constraint sets. The first column names the constraint set. In addition to the risk and tax controls used in looking at the impact of risk and taxes on the cross-sectional plots, Exhibit 8.6 lists a more detailed stepwise imposition of control matching constraints. In particular, it adds to the risk and tax controls additional controls for growth and profitability.

The next three columns in Exhibit 8.7 give high-minus-low returns. For 30 fractile portfolios in column 2, this high-minus-low value is the long-run average

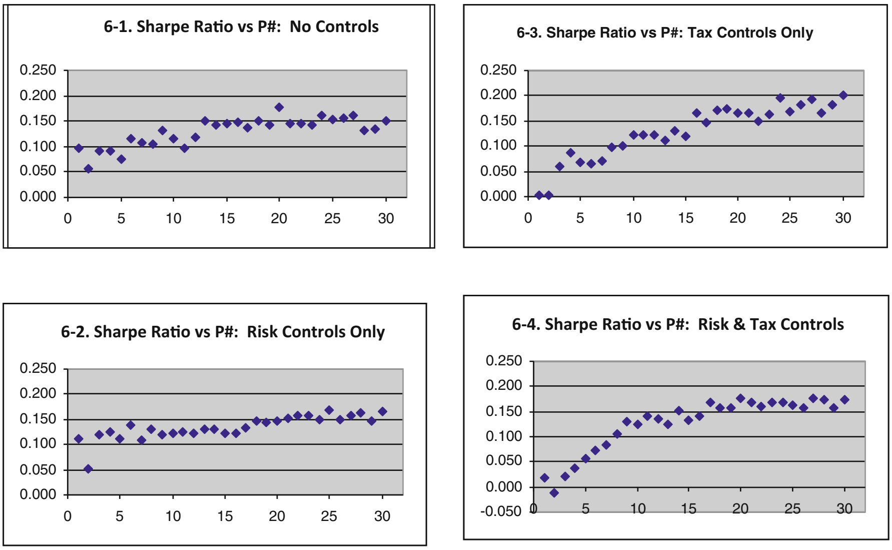


Exhibit 8.6 Sharpe ratio vs. portfolio #: risk only, tax only, and both risk and tax controls together

[30]Forecast value for mean–variance portfolio selection was established in BGMTX for 1978–1990. Guerard, Gultekin, and Stone (1997) added to the evidence of forecast value for the return forecast itself by using an endogenous APT to remove all explainable systematic return. Others have added both growth and momentum to show performance value well after the 1993 publication time.

return on portfolio 30 (the fractile portfolio with the highest forecast score) less the long-run average return on portfolio 1 (the fractile with the lowest forecast score). Column 3 is the average return for the two highest fractiles minus the average return for the two lowest fractiles. Column 4 for deciles[31] is the average of the top three fractiles minus the average of the bottom three fractiles. Since all portfolios in each cross section are matched to the ex ante values of the listed factor controls, **the high-minus-low values are the long-run realized returns on a factor-neutral arbitrage portfolio, i.e., a portfolio that is long in one or more of the top 30 fractile portfolios and short in the corresponding low score portfolios**. It is *factor neutral* in the ex ante values of each of the imposed control variables because each of the portfolios in each cross section has been matched to the sample average value of

CONTROL VARIABLES	30-tiles	Deciles	Quintiles
No Constraints	0.088	0.074	0.065
β(Beta)	0.121	0.091	0.066
S(Size)	0.086	0.075	0.064
BP(Book-to-Market)	0.076	0.068	0.055
β, S, BP	0.087	0.079	0.056
FL, EP, DP	0.196	0.176	0.137
β, S, BP, FL, EP, DP	0.211	0.165	0.128
β, S, DP, FL, Sag5, Sug3	0.239	0.187	0.147
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE	0.267	0.193	0.143
β, S, BP, FL, EP, DP, Sag5, Sug3, ROI	0.250	0.190	0.143
β, S, BP, FL, EP, DP, Sag5, Sug3, SI	0.254	0.194	0.143
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE, SI	0.276	0.198	0.145
β,S,BP,FL,EP,DP,SAG5,SUG3,ROE,SI,CP,SP	0.191	0.142	0.111

Notes:

1. For 30-tiles, the high-minus-low values for each control set are computed by taking the difference between the average realized return for portfolio 30 and for portfolio 1.
2. For deciles, the high-minus-low value is the difference between the average realized returns for the top 3 portfolios minus the average realized return for the bottom 3 portfolios.
3. For quintiles, the high-minus-low value is the difference between the average realized return in the top six portfolios minus the average realized return for the bottom six portfolios.

Exhibit 8.7 High-minus-low values for 1968–2004 average returns: how imposing controls changes the extreme high and low returns

[31]With 30 fractile portfolios in the cross sections of conditional return response observations, the difference for the top three returns combined and the bottom three portfolios combined represents a high-minus-low return for the top and bottom deciles of the cross section.

the imposed controls. Therefore, **a long–short combination of any two portfolios has zero ex ante exposure to the imposed controls**.

The high-minus-low values are the annualized average of 458 monthly values. Thus, they indicate the economic significance of the composite value score before any transaction costs for a naive factor-neutral portfolio strategy.

The term "naïve" refers to the fact that these portfolios are formed on the basis of return forecast data alone without using any information about variances, covariances, or higher moments such as skewness. Given that past values of both variance and covariance are fairly good predictions of month-ahead values, use of the value-focused return scores with mean–variance optimization should always produce superior market-neutral hedge portfolios in terms of Sharpe ratios. **The Sharpe ratios reported for these three market-neutral portfolios are lower bounds on the mean–variance optimized factor-neutral portfolios**.

The high-minus-low ranges and associated Sharpe ratios both exhibit a strong dependency on the imposed controls. The return cross section with no controls has a range of 8.8%, the Fama-French three factor control set has a range of 8.7%. In contrast to the is very small change from imposing the three risk controls, imposing the three tax controls results in a high-minus-low range of 19.6%. Imposing the three risk controls and the three tax controls together further increases the 30-fractile high-minus-low range to 21.1%.

Adding growth controls increases the hml to more than 24 %, triples the range for the unranked cross section, and more than doubles the range when the only control variables are the conventional Fama–French three-factor risk instruments. Adding sales intensity and profitability controls further increases the range and improves the Sharpe ratios.

8.8.10 Estimates of the Dependence of the Return and SD Cross Sections on the Return Forecast

Exhibits 8.8, 8.9, and 8.10 summarize regression tests of the ability of the return forecast score to explain the long-run realized cross sections of average returns, Sharpe ratios, standard deviations, and skewness coefficients.

For the stepwise imposition of a series of control constraints, the first table in Exhibit 8.8 summarizes linear regressions of the long-run 456 month average value of realized risky return on return forecast score for a series of progressively more complete sets of control variables.

Control Variables	$R_p = C_0 + C_1(FS_p) + \varepsilon_p$			
	slope	t	R^2	p-value
No Constraints	.079	7.23	.651	<.0001
β(Beta)	.086	7.47	.666	<.0001
S(Size)	.080	7.91	.691	<.0001
BP(Book-to-Market)	.071	7.77	.683	<.0001
β, S, BP	.079	6.44	.597	<.0001
FL, EP, DP	.181	17.81	.919	<.0001
β, S, BP, FL, EP, DP	.192	18.97	.928	<.0001
β, S, DP, FL, Sag5, Sug3	.206	16.4	.906	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE	.217	14.57	.884	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROI	.215	16.44	.906	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, SI	.216	15.66	.898	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE, SI	.220	14.24	.879	<.0001
β,S,BP,FL,EP,DP,SAG5,SUG3,ROE,SI,CP,SP	.186	16.56	.907	<.0001

Exhibit 8.8 The changing ability of forecast score to explain realized returns and Sharpe ratios

All of the regressions in Exhibit 8.8 have very high R-squared values, large and significant *t*-values, and *p*-values less than 0.0001. Given that all of the regressions have *p*-values less than 0.0001, the change in the *t*-values for the coefficient on return forecast score and the change in *R*-squared are the best indicators of the effect on the cross section of imposing additional control constraints, especially in terms of the extent to which we are obtaining a return forecast dependency that is better isolated from the effect of nonforecast variables. Thus, we focus most of our attention here on the changes in the *t*-values and *R*-squared values as we impose different sets of control constraints.

The cross section of realized returns with no controls and with the three risk controls imposed individually and in combination results in very modest changes in both the estimated slope and the associated *t*-value. For instance, the imposition of the three risk controls together produces a *t*-value on the slope coefficient of 0.597 with an associated *t*-value of 6.44 compared to the no-control case of a slope coefficient of 0.079 and a t-value of 7.23.

The most significant structural feature is the jump in *R*-squared values and *t*-values when we impose the tax controls alone or impose the tax controls along with the three systematic risk factors. Imposing the three tax controls alone produces a t-value of 17.81. The control set with both systematic risk and tax controls has a t-value of 18.97, a clearly significant increase in the slope estimate and its significance.

These results are a surprise! First, the negligible impact of the three Fama–French risk controls on the long-run cross section of realized returns means that apparent return potential is not a systematic risk effect in disguise, at least in terms of the three Fama–French risk variables. The major surprise is the large and very significant tax effect. **Most surprising is the direction of the tax effect on the assessment of performance potential**. In noting the issue of distortion from regularly recurring systematic tax effects, the concern was that the very high correlation of the four current value ratios with dividend yield could mean that apparent forecast potential could be a dividend yield tilt in disguise rather than finding truly misvalued stocks that would produce superior returns as the market recognized the undervaluation.

Also surprising is the fact that imposing the combination of risk and tax controls means that all portfolios in the cross section have the same portfolio average values of BP and EP, two of the eight variables weighted in the return forecast prediction. In effect, any contribution of these two variables to realized return performance is suppressed for the combination of risk and tax controls. The apparently large and significant performance is from the other six variables. It is pertinent to note that a benefit of the control approach is a straightforward assessment of the relative contribution of one or more of the forecast variables in a multivariate forecasting model. Contrary to much empirical evidence on the value of BP and EP for explaining the cross section of realized returns, it appears that for the cross section based on the BGMTX forecast model, neither BP nor EP are an important part of the very significant performance potential indicated by the plots, the high-minus-low returns, and especially the very significant slope for the linear regression fit. This analysis indicates that most of the forecast potential must arise from the other six variables.

In order to gain more insight on forecast potential, it is useful to add to the risk and tax controls additional controls for growth and profitability and also to assess the impact on the return cross section of suppressing other model variables. The next control sets summarized in Exhibit 8.8 add to the Fama–French risk controls and the set of tax controls two growth controls (5-year sales growth, 3-year sustainable growth) plus a profitability control (either ROE, ROI, or sales intensity SI). In all three cases the estimated slope coefficient increases to more than 0.20, but both the t-value and R^2 decrease slightly. The increased slope coefficient with poorer fit makes sense if one also considers the high-minus-low data in Exhibit 8.7. The effect of adding growth and profitability controls is to increase primarily the very high returns and to decrease the very low returns and to thereby increase the departure from linearity. Hence, there is more range and more performance potential for the very high predicted returns but a departure from linearity and a poorer linear fit.

The final control set adds controls for two more forecast variables: CP and SP. Hence, for this control set there is no cross-sectional impact from any of the four current value ratios: BP, EP, CP, and SP. The large significant slope coefficient is attributable solely to the four relative value ratios. **If we use the estimated slope coefficient as an indicator of overall ability of the return forecast to predict risk-controlled, tax-controlled, growth–profitability controlled realized return predictions, the suppression of CP and SP indicates that the relative value ratios are responsible for about 80 % of the apparent return forecast potential, CP and SP contribute about 20 % of the apparent return forecast potential, and BP and EP seem to contribute very little to the apparent return forecast potential.**

8.8.11 The Cross Sections of Realized Standard Deviations for Different Combinations of Controls

Exhibit 8.9 summarizes regressions of the realized cross-time standard deviations on return forecast score for different combinations of controls.

The coefficient on the linear term C_1 is insignificant until tax controls (FL, EP, and DP) are imposed. The large jump in the *t*-value with the imposition of tax controls alone or in combination with other variables again indicates that controlling for tax effects is critical to isolate return forecast performance from other distorting return factors. For the cross section of realized standard deviations, it appears systematic tax effects are the most pertinent set of control variables rather than the usual systematic risk variables.

In all of the cross-sectional plots summarizing the dependence of realized standard deviation on portfolio number such as in Exhibit 8.5, the dependence of realized standard deviation on portfolio number is clearly nonlinear and nonmonotonic. In particular, the standard deviations for low portfolio numbers (low forecast scores) and for high portfolio numbers (high forecast scores) were all substantially greater than the standard deviations for the middle of the cross section. Visual inspection of the cross-sectional plots suggests a quadratic dependency. For this reason, the regressions summarized in Exhibit 8.9 designed to assess the cross-sectional dependence of realized standard deviations on return forecast score include a quadratic term as well as a linear term. Because the concern is assessing the impact of below average and above average forecast scores relative to the average, the quadratic dependency is expressed as a squared deviation of return forecast score

Control Variables	$SD_p = C_0 + C_1(FS_p) + C_2(FS_p - mean(FS_p))^2 + \varepsilon_p$					
	C_1	t	C_2	t	Adj R^2	p-value
No Constraints	0.07	2.30	0.02	18.21	0.920	<.0001
β(Beta)	0.11	3.74	0.01	12.45	0.852	<.0001
S(Size)	0.07	2.31	0.02	18.36	0.921	<.0001
BP(Book-to-Market)	0.11	3.89	0.02	22.24	0.946	<.0001
β, S, BP	0.14	5.50	0.02	15.46	0.902	<.0001
FL, EP, DP	0.18	6.46	0.02	17.32	0.928	<.0001
β, S, BP, FL, EP, DP	0.20	9.98	0.01	16.57	0.934	<.0001
β, S, DP, FL, Sag5, Sug3	0.18	7.18	0.01	13.27	0.895	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE	0.21	9.24	0.01	14.80	0.918	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROI	0.20	10.20	0.01	16.20	0.931	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, SI	0.21	9.44	0.01	15.04	0.921	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE, SI	0.21	9.62	0.01	15.11	0.922	<.0001
β,S,BP,FL,EP,DP,SAG5,SUG3,ROE,SI,CP,SP	0.17	11.06	0.01	17.13	0.936	<.0001

Exhibit 8.9 The ability of forecast score to explain the cross section of standard deviations

from the average forecast score. In the regressions in Exhibit 8.9, the t-values for the quadratic coefficient C_2 are much larger and thus much more significant than the t-values for the linear coefficient C_1. The very high adjusted R-square values for all of the regressions in Exhibit 8.9 strongly indicate that the combination of a linear and quadratic dependency explains most of the cross-sectional variation in the realized standard deviations. The much higher significance for the quadratic term is confirmation of the importance of the nonlinear, nonmonotonic apparently near-quadratic dependency suggested by the cross-sectional plots such as Exhibit 8.5.

8.8.12 The Cross Section of Realized Skewness Coefficients

Exhibit 8.10 is a cross-sectional regression on the long-run realized skewness as measured by the skewness coefficient.

The skewness is significant even with no control constraints. This jump in the *t*-value from imposing the beta control alone suggests that controlling for market movements by means of the beta control increases the isolation of nonsystematic skewness from any market skewness. Interestingly, neither the size control alone nor the book-to-market control alone significantly changes the skewness. However, the three systematic risk controls together increase the increase skewness.

In contrast to the regressions for realized standard deviation, tax controls alone do not seem to help isolate skewness effects. However, the three risk controls plus the three tax controls together do increase the coefficients and associated *t*-values.

Control Variables	$SD_p = C_0 + C_1(FS_p) + C_2(FS_p - mean(FS_p))^2 + \varepsilon_p$							
	C_0	t	C_1	t	C_2	t	Adj R^2	p-value
No Constraints	-1.272	-14.36	0.014	10.7	0.001	13.67	0.911	<.0001
β(Beta)	-1.252	-14.06	0.017	12.13	0.000	8.55	0.883	<.0001
S(Size)	-1.274	-13.42	0.014	9.95	0.001	12.84	0.900	<.0001
BP(Book-to-Market)	-1.375	-15.84	0.016	11.7	0.001	14.68	0.923	<.0001
β, S, BP	-1.340	-21.18	0.018	17.84	0.000	12.14	0.941	<.0001
FL, EP, DP	-1.164	-9.15	0.013	5.92	0.001	9.16	0.822	<.0001
β, S, BP, FL, EP, DP	-1.365	-17.42	0.019	14.07	0.000	7.93	0.907	<.0001
β, S, DP, FL, Sag5, Sug3	-1.347	-21.04	0.018	17.12	0.001	11.15	0.940	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE	-1.408	-18.3	0.020	15.07	0.000	8.2	0.916	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROI	-1.411	-23.11	0.020	19.17	0.000	10	0.945	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, SI	-1.427	-19.23	0.020	15.88	0.001	9.24	0.926	<.0001
β, S, BP, FL, EP, DP, Sag5, Sug3, ROE, SI	-1.418	-21.82	0.020	18.03	0.000	10.02	0.940	<.0001
β,S,BP,FL,EP,DP,SAG5,SUG3,ROE,SI,CP,SP	-1.150	-18.14	0.016	14.51	0.000	5.89	0.896	<.0001

Exhibit 8.10 The ability of forecast score to explain the cross section of skewness coefficients

Adding our two growth controls and other company-specific controls does increase significantly the ability of return forecast score to explain the cross-sectional skewness coefficient. We conclude that isolating nonsystematic value-related skewness associated with the illustrative value-focused return forecasting model requires that we isolate these value-related return and risk effects from growth in particular.

In the final set of controls in Exhibit 8.10, we add two more controls for two of the forecast model variables, namely, CP and SP. Given that the control set already contains the BP and EP ratios, the net effect of this final control set is to remove from the cross section the contribution to realized returns of all four current value ratios. Thus, this final set thus measures the skewness response for just the relative value ratios. It is interesting that the *t*-value is the greatest for this set of control constraints. **This high *t*-value for the response of realized skewness to the relative value subset of the eight forecast variables is strong evidence that the significant cross-sectional dependency of skewness on return forecast score is primarily from the relative value ratios rather than the current value ratios**.

Since the relative value ratios measure attractiveness for a company relative to its own past value ratios, this result suggests, or at least is consistent with, company values returning to a moving mean. In other words, this result indicates that when current value ratios are well below their 5-year average value, there is an apparently strong likelihood that the stock price will increase in order to return the company to its recent average valuation. **This result is consistent with the value analyst use of relative value ratios to find stocks with limited downside risk and upside potential**. It also suggest turning point performance, rather than trend (momentum) as the primary source of unpriced, unsystematic skewness.

8.9 Further Research

Having good return forecasts is the primary requirement for successful active portfolio management. Given one or more return forecasting alternatives, the central problem for making decisions about return forecast centers on the ability to conduct high power, high efficiency backtest assessments. The focus here has been to illustrate the use of a mathematical assignment program to optimize the construction of control-matched cross-sections of rank-ordered return forecasts. The central requirement for high-quality return forecast assessments is the ability to isolate the impact of the return forecast from the impact of other return variables. The control matching framework is an alternative to the use of multivariate regression to evaluate return forecasts.

The BGMTX forecast model has been used as an illustrative multivariate return forecast. There are several ways to extend this forecast assessment. A very important question for forecast performance is cross-time consistency. The majority of this chapter has looked at the very long 1967-2004 for time period. Stone and Gerard

(2010) report comparable performance results for this United States sample for four subperiods of 1967–2004.

Additional performance concerns are relative ability to perform in up-markets versus down-markets as well as performance in other markets or market subsectors. Stone and Gerard (2011) use control matching to evaluate performance for Japan including subperiods of net market decline. The Japan assessment shows that the BGMTX forecast model rank orders return performance over extended net market declines in Japan, for instance excellent relative rank ordering for control matched cross-sections when every fractile portfolio except the highest ranked has a negative return.

Tracking error pertains to cross-time consistency. Given a sample of control matched cross-sections for a forecast time period such as the 456 monthly control matched cross-sections used in this illustration of the methodology, a researcher can evaluate tracking error by looking at moving averages, e.g., 12-month, 24-month, and 60-month rolling averages to obtain an indication of one-year, two-year, and five-year performance consistency risk. Observing moving average performance provides insights to not only tracking consistency but also relative performance in net up markets, net down markets, and over market reversals. Relative performance in up, down, and reversal markets provides insight on the extent to which a particular forecast is momentum focused (trend extrapolating) versus turning-point focused. For the BGMTX forecast model used in this illustration of control matching, one-year moving averages indicate that much of the performance value occurred over marker reversals, both transition from a bear to bull market and also transition from a bull market to a bear market. This reversal performance is consistent with the fact that suppressing the effect of the four value ratios indicated that roughly 80% of the high-minus-low performance summarized in Exhibit 8.7 was attributable to the four relative value (return reversal) variables in the BGMTX forecast model.

The key idea implicit in the preceding discussion is many alternative performance assessment insights beyond the statistical tests used in this chapter. Having a time series of control matched forecast cross sections allows a researcher to investigate not only statistical return performance but also cross-sectional risk (tracking error, standard deviations, skewness, etc).

Looking at performance value without noise from other return variables lets a researcher experiment with forecast performance without distortion from other return impact variables. Moreover, looking at performance for alternative control sets provides important information on interaction with other return impact variables. For instance, the fact that the two growth controls had an impact on the high-minus-low performance even after controlling for both risk and tax effects suggests a nonsystematic growth effect and therefore potential value to adding one or more growth variables to the BGMTX forecast model.

8.10 Conclusions

The BGMTX forecast model shows significant return performance potential as established in other studies. Compared to studies that do not use matched controls, the use of risk, tax, growth, and profitability controls provides additional information.

1. Both the return and the realized standard deviations cross sections are nonlinear. The cross section of realized standard deviations is a relatively small range compared to the range in the cross section of realized returns. The cross section of realize standard deviations is not only nonlinear but highly nonmonotonic.
2. The distribution of realized returns about the average value exhibits skewness, negative skewness for low return forecasts, very little for the middle of the distribution, and very large significant skewness for the highest return forecasts.
3. The three risk control variables tend to smooth the cross sections of realized returns; however, risk variables appear not to have a significant effect on the long-run cross section of realized returns. These risk variables are not a source of systematic performance bias.
4. The three risk controls in combination tend to smooth the cross sections of realized standard deviations. More importantly, the three risk controls together reduce the range of realized standard deviations.
5. Tax effects are very significant for the illustrative BGMTX forecast model. Contrary to the hypothesis of apparent return potential being a tax tilt, imposing the three tax controls to eliminate cross-sectional variation in the dividend–gain mix significantly increases the slope and range of the realized return cross section and moderately reduces realized standard deviation.
6. The power optimizing imposition of a combination of risk and tax controls significantly increases statistical efficiency relative to the uncontrolled cross section. Adding growth and profitability controls adds additional value in assessing return potential.
7. Suppressing the four current value ratios in the cross section of realized returns reduces the slope of the realized return cross section only modestly, about 20 %. The relative value ratios are the major source of realized return potential and the significant positive skewness for the high-return forecast part of the cross section.

The main methodology benefit for full sample assessments of performance potential of a return forecast is the control matching framework itself. Using a power optimizing reassignment program like that formulated in Sect. 8.5 makes it possible to eliminate return performance distortion from other return impacting variables without having to make any assumptions about the distribution or form of the functional dependency for any of the control variables. One simply has to assume a possible dependency on the return variables. Imposing the control constraint eliminates any portfolio-to-portfolio impact from the matched controls. The distribution-free, specification-free attribute avoids functional form specification errors and the estimation limitations associated with distributional assumptions.

Compared to using a multivariate regression to estimate jointly the sensitivity of realized returns to the forecast and also to estimate an assumed dependency for all the other possible return impact variables for risk, taxes, growth, profitability, etc., the use of controls to suppress cross-sectional variation in the other variable dependencies means only the need to estimate a univariate dependency on the return forecast under the assumption that any effect of other return impact variables has been suppressed. Concentrating all sample data on a conditional univariate dependency rather than estimating a multivariate dependency means much more efficient, more powerful extraction of sample information relative to making restrictive assumptions to estimate multiple dependencies when the concern is a single well-isolated conditional dependency.

The use of tax controls illustrates the benefit of avoiding functional specification to remove a potential source of realized return distortion. By simply making every portfolio in the cross section have the same dividend yield, dividend payout ratio, and same benefit from the debt tax shield, it is not necessary to estimate time-varying marginal tax rates for dividends and gains. Similarly, using FL as a financial control avoided the need to assess the value of the debt tax shield and the simultaneous need to correct for distortion from other possible but not known valuation effects of financial leverage.

In addition to the benefits of concentrating data on a univariate dependency and avoiding distribution and specification assumptions, the control matching framework completely eliminates bias/distortion from covariability effects between the return forecast and any of the control variables. Complete elimination of covariability contamination is a significant power benefit!

Another efficiency/power design concern is the number of portfolios in the cross section. The efficiency benefit of grouping observations is reduced measurement error and possibly reduced specification and omitted variable error. Grouping observations loses power by reducing the number of sample observations. The point made in this paper is the need to explicitly recognize the relative benefit of reduced measurement error versus loss of power from reduction in the number of sample observations and the associated loss of information from using averages to represent a collection of observations. In the control matching framework, it is possible to consolidate control-matched fractiles and preserve the control matching. This consolidation is illustrated in Sect. 8.8.9 assessing the high-minus-low realized return differences for 30 tiles and the consolidation to 15 tiles and deciles.

Return forecasting is very much an art using knowledge of valuation and statistics. Assessing return forecast performance potential is also an art. The matched control framework aided by an optimal reassignment algorithm is a decision support framework for exploring return performance potential that avoids many limitations of multivariate regression assessments including especially collinearity distortion and the restrictions of distributional and functional form assumptions.

Appendices

Appendix 8.1. Rank-ordered portfolio data: no controls

P#	FS	Rtn%	SD%	Skew	S Ratio
1	2.98	10.29	8.96	0.802	0.096
2	3.90	6.22	9.11	0.279	0.057
3	6.99	8.84	7.96	0.065	0.093
4	11.70	7.72	7.01	−0.231	0.092
5	14.38	6.10	6.81	−0.300	0.075
6	18.02	8.89	6.41	−0.543	0.116
7	21.51	7.94	6.13	−0.509	0.108
8	24.98	7.75	6.09	−0.519	0.106
9	28.05	9.18	5.76	−0.682	0.133
10	31.50	7.85	5.68	−0.593	0.115
11	34.84	6.48	5.54	−0.714	0.097
12	37.98	7.96	5.56	−0.606	0.119
13	41.04	9.92	5.48	−0.591	0.151
14	44.48	9.20	5.43	−0.528	0.141
15	47.98	9.39	5.39	−0.503	0.145
16	51.03	9.42	5.35	−0.384	0.147
17	54.45	8.79	5.35	−0.417	0.137
18	57.89	9.93	5.54	−0.213	0.149
19	60.96	9.48	5.50	−0.209	0.144
20	64.44	11.74	5.56	−0.241	0.176
21	67.87	9.97	5.69	−0.077	0.146
22	70.89	10.04	5.78	0.034	0.145
23	73.96	10.36	6.08	0.462	0.142
24	77.43	12.39	6.35	0.287	0.163
25	80.87	11.95	6.49	0.530	0.154
26	84.41	12.93	6.95	0.421	0.155
27	87.06	14.29	7.35	1.061	0.162
28	91.08	13.03	8.26	0.993	0.131
29	94.29	15.50	9.57	1.271	0.135
30	96.59	19.12	10.62	2.308	0.150

Appendix 8.2. Rank-ordered portfolio data: only a beta control

P#	FS	Rtn%	SD%	Skew	S Ratio
1	4.21	7.77	7.36	−0.248	0.088
2	5.20	5.74	7.59	0.080	0.063
3	8.54	9.37	7.14	−0.447	0.109
4	13.52	7.72	6.22	−0.617	0.103
5	16.19	7.64	6.40	−0.636	0.099
6	19.69	9.61	5.83	−0.565	0.137
7	23.12	7.73	5.98	−0.612	0.108
8	26.27	7.60	5.81	−0.477	0.109
9	29.14	7.56	5.78	−0.507	0.109
10	32.34	9.83	5.68	−0.629	0.144
11	35.49	7.94	5.56	−0.400	0.119
12	38.51	7.73	5.77	−0.541	0.112
13	41.44	7.68	5.61	−0.567	0.114
14	44.72	8.70	5.64	−0.402	0.128
15	48.02	8.08	5.54	−0.484	0.121
16	50.97	9.39	5.65	−0.404	0.139
17	54.24	8.85	5.59	−0.290	0.132
18	57.51	9.04	5.76	−0.034	0.131
19	60.47	9.17	5.75	0.197	0.133
20	63.79	11.05	5.84	−0.012	0.158
21	67.07	10.47	5.78	−0.180	0.151
22	69.96	10.37	5.92	0.149	0.146
23	72.92	10.89	5.85	−0.010	0.155
24	76.15	12.23	6.09	0.042	0.167
25	79.65	12.76	6.07	0.287	0.175
26	83.18	11.10	6.50	0.632	0.142
27	85.85	13.27	6.78	0.247	0.163
28	89.97	15.10	7.65	1.263	0.164
29	93.48	15.11	8.68	1.483	0.145
30	95.94	19.89	9.55	1.585	0.173

Appendix 8.3. Rank-ordered portfolio data: only a size control

P#	FS	Rtn%	SD%	Skew	S Ratio
1	3.28	10.31	8.90	0.914	0.097
2	4.22	6.77	9.08	0.159	0.062
3	7.31	7.73	7.93	0.137	0.081
4	12.00	7.91	7.03	−0.313	0.094
5	14.65	6.35	6.85	−0.200	0.077
6	18.28	8.70	6.33	−0.556	0.115
7	21.86	8.06	6.06	−0.621	0.111
8	25.20	7.79	6.10	−0.538	0.106
9	28.29	9.07	5.84	−0.602	0.130
10	31.67	7.09	5.61	−0.704	0.105
11	34.97	6.80	5.54	−0.640	0.102
12	38.11	8.49	5.54	−0.537	0.128
13	41.16	9.49	5.47	−0.581	0.145
14	44.51	9.05	5.35	−0.604	0.141
15	47.99	9.34	5.44	−0.413	0.143
16	51.09	9.30	5.33	−0.375	0.145
17	54.40	8.84	5.36	−0.478	0.137
18	57.82	10.07	5.58	−0.279	0.150
19	60.90	9.79	5.42	−0.198	0.150
20	64.26	11.35	5.51	−0.196	0.172
21	67.67	10.17	5.62	−0.035	0.151
22	70.68	10.91	5.95	0.266	0.153
23	73.75	10.31	5.96	0.189	0.144
24	77.07	12.86	6.31	0.340	0.170
25	80.63	11.73	6.52	0.322	0.150
26	84.15	12.78	6.96	0.514	0.153
27	86.80	14.43	7.29	1.045	0.165
28	90.80	13.24	8.21	1.017	0.134
29	94.01	15.06	9.58	1.261	0.131
30	96.28	18.87	10.55	2.329	0.149

Appendix 8.4. Rank-ordered portfolio data: only a BP control

P#	FS	Rtn%	SD%	Skew	S Ratio
1	4.03	10.27	8.83	0.805	0.097
2	5.47	7.47	8.89	0.276	0.070
3	9.10	7.66	7.64	−0.064	0.083
4	14.47	7.46	6.83	−0.460	0.091
5	17.26	8.51	6.71	−0.495	0.106
6	21.01	8.60	6.45	−0.292	0.111
7	24.59	9.23	5.94	−0.657	0.129
8	27.75	7.76	5.88	−0.745	0.110
9	30.61	8.45	5.87	−0.772	0.120
10	33.74	8.09	5.67	−0.515	0.119
11	36.75	7.70	5.38	−0.824	0.119
12	39.56	10.29	5.55	−0.547	0.154
13	42.27	9.68	5.51	−0.690	0.146
14	45.27	8.96	5.38	−0.569	0.139
15	48.26	8.93	5.25	−0.594	0.142
16	50.89	9.74	5.35	−0.425	0.152
17	53.84	9.24	5.33	−0.381	0.144
18	56.82	10.67	5.37	−0.310	0.166
19	59.50	10.67	5.51	−0.077	0.161
20	62.58	10.70	5.49	−0.173	0.162
21	65.68	9.64	5.72	−0.105	0.140
22	68.44	10.50	5.83	−0.038	0.150
23	71.36	11.25	6.02	0.177	0.156
24	74.60	10.88	6.30	0.220	0.144
25	78.16	11.75	6.70	0.565	0.146
26	81.86	12.03	6.92	0.373	0.145
27	84.69	13.33	7.15	0.707	0.155
28	89.21	12.84	8.44	1.469	0.127
29	92.95	15.15	9.62	1.342	0.131
30	95.83	17.85	10.65	2.186	0.140

Appendix 8.5. Rank-ordered portfolio data: risk controls only

P#	FS	Rtn%	SD%	Skew	S Ratio
1	4.21	7.77	7.36	−0.248	0.088
2	5.20	5.74	7.59	0.080	0.063
3	8.54	9.37	7.14	−0.447	0.109
4	13.52	7.72	6.22	−0.617	0.103
5	16.19	7.64	6.40	−0.636	0.099
6	19.69	9.61	5.83	−0.565	0.137
7	23.12	7.73	5.98	−0.612	0.108
8	26.27	7.60	5.81	−0.477	0.109
9	29.14	7.56	5.78	−0.507	0.109
10	32.34	9.83	5.68	−0.629	0.144
11	35.49	7.94	5.56	−0.400	0.119
12	38.51	7.73	5.77	−0.541	0.112
13	41.44	7.68	5.61	−0.567	0.114
14	44.72	8.70	5.64	−0.402	0.128
15	48.02	8.08	5.54	−0.484	0.121
16	50.97	9.39	5.65	−0.404	0.139
17	54.24	8.85	5.59	−0.290	0.132
18	57.51	9.04	5.76	−0.034	0.131
19	60.47	9.17	5.75	0.197	0.133
20	63.79	11.05	5.84	−0.012	0.158
21	67.07	10.47	5.78	−0.180	0.151
22	69.96	10.37	5.92	0.149	0.146
23	72.92	10.89	5.85	−0.010	0.155
24	76.15	12.23	6.09	0.042	0.167
25	79.65	12.76	6.07	0.287	0.175
26	83.18	11.10	6.50	0.632	0.142
27	85.85	13.27	6.78	0.247	0.163
28	89.97	15.10	7.65	1.263	0.164
29	93.48	15.11	8.68	1.483	0.145
30	95.94	19.89	9.55	1.585	0.173

Appendix 8.6. Rank-ordered portfolio data: tax controls only

P#	FS	Rtn%	SD%	Skew	S Ratio
1	10.42	0.30	6.50	−0.420	0.004
2	11.42	0.15	6.45	−0.393	0.002
3	15.46	4.50	6.35	−0.600	0.059
4	20.52	6.09	5.89	−0.737	0.086
5	22.84	4.93	5.98	−0.495	0.069
6	25.76	4.61	5.87	−0.728	0.065
7	28.62	4.90	5.67	−0.642	0.072
8	31.02	6.47	5.57	−0.729	0.097
9	33.27	6.75	5.59	−0.490	0.100
10	35.77	8.13	5.60	−0.638	0.121
11	38.14	7.96	5.47	−0.645	0.121
12	40.46	8.12	5.49	−0.587	0.123
13	42.74	7.33	5.47	−0.607	0.112
14	45.30	8.78	5.59	−0.522	0.131
15	47.90	8.26	5.70	−0.377	0.121
16	50.26	10.91	5.51	−0.407	0.165
17	52.94	9.86	5.60	−0.277	0.147
18	55.63	11.84	5.77	−0.077	0.171
19	58.13	12.02	5.72	−0.269	0.175
20	60.96	11.56	5.83	−0.010	0.165
21	63.80	11.84	5.92	0.047	0.167
22	66.32	10.84	6.02	0.019	0.150
23	68.98	11.79	6.07	0.147	0.162
24	71.98	14.23	6.05	0.069	0.196
25	75.37	12.79	6.38	0.144	0.167
26	78.92	13.99	6.39	0.121	0.182
27	81.66	16.00	6.90	0.604	0.193
28	86.26	14.75	7.38	0.515	0.166
29	90.47	18.25	8.38	0.987	0.181
30	93.65	21.40	8.92	1.789	0.200

Appendix 8.7. Rank-ordered portfolio data: risk and tax controls

P#	FS	Rtn%	SD%	Skew	S Ratio
1	10.42	0.30	6.50	−0.420	0.004
2	11.42	0.15	6.45	−0.393	0.002
3	15.46	4.50	6.35	−0.600	0.059
4	20.52	6.09	5.89	−0.737	0.086
5	22.84	4.93	5.98	−0.495	0.069
6	25.76	4.61	5.87	−0.728	0.065
7	28.62	4.90	5.67	−0.642	0.072
8	31.02	6.47	5.57	−0.729	0.097
9	33.27	6.75	5.59	−0.490	0.100
10	35.77	8.13	5.60	−0.638	0.121
11	38.14	7.96	5.47	−0.645	0.121
12	40.46	8.12	5.49	−0.587	0.123
13	42.74	7.33	5.47	−0.607	0.112
14	45.30	8.78	5.59	−0.522	0.131
15	47.90	8.26	5.70	−0.377	0.121
16	50.26	10.91	5.51	−0.407	0.165
17	52.94	9.86	5.60	−0.277	0.147
18	55.63	11.84	5.77	−0.077	0.171
19	58.13	12.02	5.72	−0.269	0.175
20	60.96	11.56	5.83	−0.010	0.165
21	63.80	11.84	5.92	0.047	0.167
22	66.32	10.84	6.02	0.019	0.150
23	68.98	11.79	6.07	0.147	0.162
24	71.98	14.23	6.05	0.069	0.196
25	75.37	12.79	6.38	0.144	0.167
26	78.92	13.99	6.39	0.121	0.182
27	81.66	16.00	6.90	0.604	0.193
28	86.26	14.75	7.38	0.515	0.166
29	90.47	18.25	8.38	0.987	0.181
30	93.65	21.40	8.92	1.789	0.200

References

Basu, S. (1977). Investment performance of common stocks in relations to their price earnings ratios: A test of market efficiency. *Journal of Finance, 32*, 663–682.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the theory of capital markets* (pp. 79–122). New York: Praeger Publishers.

Blume, M. (1980). Stock return and dividend yield: Some more evidence. *Review of Economics and Statistics, 62*, 567–577.

Box, G. E. P., Hunter, J. S., & Hunter, W. G. (2005). *Statistics for experimenters*. New York: Wiley Interscience.

Box, G. E. P. (1954). The exploration and exploitation of response surfaces. *Biometrics, 10*, 16–60.

Box, G. E. P., & Draper, N. R. (1987). *Empirical model building and response surfaces*. New York: Wiley Interscience.

Box, G. E. P., & Draper, N. R. (1959). A basis for the selection of a response surface design. *Journal of the American Statistical Association, 54*, 622–654.

Brennen, M. (1970). Taxes, market valuation and financial policy. *National Tax Journal, 23*, 417–429.

Brown, S. J. (2008). Elusive return predictability: Discussion. *International Journal of Forecasting, 24*, 19–21.

Conner, G., & Korajczak, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics, 21*, 255–289.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–25.

Fama, E. F., & French, K. R. (1992). Cross-sectional variation in expected stock returns. *Journal of Finance, 47*, 427–465.

Fama, E. F., & French, K. R. (2008a). Dissecting anomalies. *Journal of Finance, 63*, 1653–1678.

Fama, E. F., & French, K. R. (2008b). Average returns, B/M, and share issues. *Journal of Finance, 63*, 2971–2995.

Fama, E. F., & MacBeth, J. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*, 607–636.

Graham, B., & Dodd, D. L. (1934). *Security analysis*. New York: McGraw-Hill Book Company.

Griliches, Z. (1986). Economic data issues. In Z. Griliches & M. Intrilagator (Eds.), *Handbook of econometrics*. Amsterdam: North Holland.

Grinold, R. C., & Kahn, R. N. (2000). *Active portfolio management* (2nd ed.). New York: McGraw-Hill.

Grunfeld, Y., & Griliches, Z. (1960). Is aggregation necessarily bad? *Review of Economics and Statistics, 42*, 1–13.

Guerard, J. B., Jr., Gultekin, M. G., & Stone, B. K. (1997). The role of fundamental data and analysts' earnings breadth, forecasts, and revisions in the creation of efficient portfolios. In A. Chen (Ed.), *Research in finance 15*, 69–91.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance, 23*, 389–416.

Khuri, A. I., & Cornell, J. A. (1996). *Response surfaces: Designs and analysis*. New York: Marcel Dekker, Inc.

Lease, R. C., John, K., Kalay, A., Loewenstein, U., & Sarig, O. H. (2000). *Dividend policy: Its impact on firm value*. Boston: Harvard Business School Press.

Markowitz, H. M. (1987). *Mean-variance analysis in portfolio choice and capital markets*. Oxford: Basil Blackwell.

Markowitz, H. M. (1991). Foundations of portfolio theory. *Journal of Finance, 46*(2), 469–477.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investment*. New York: John Wiley & Sons. Cowles Foundation monograph no. 16.

Markowitz, H. M., & Xu, G. (1994). Data mining corrections. *Journal of Portfolio Management, 21*, 60–69.

Menchero, J., Morozov, A., & Shepard, P. (2010). Global equity risk modeling. In J. B. Guerard Jr. (Ed.), *Chapter 15, Handbook of portfolio construction* (pp. 439–480). New York: Springer.

Michaud, R. O. (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, *45*, #1 (January/February 1989).

Michaud, R. O. (1998). *Efficient asset management*. Boston: Harvard Business School Press.

Myers, R. H., Khuri, A. I., & Cornell, W. H., Jr. (1989). Response surface methodology: 1966–1988.*Technometrics, 31*, 137–157.

Peterson, P., Peterson, D., & Ang, J. (1985). Direct evidence on the marginal rate of taxation on dividend income. *Journal of Financial Economics, 14*, 267–282.

Pilotte, E. A. (2003). Capital gains, dividend yields, and expected inflation. *Journal of Finance, 58*, 447–466.

Rosenberg, B. (1974). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis, 9*, 263–274.

Rosenberg, B., & Rudd, A. (1982). Factor-related and specific returns of common stocks: Serial correlation and market inefficiency. *Journal of Finance, 37*, 543–554.

Rosenberg, B., & Marathe, V. (1979). Tests of capital asset pricing hypotheses. In H. Levy (Ed.), *Research in finance 1*. Bingley, UK: Emerald Publishing Group Limited.

Rosenberg, B., & Rudd, A. (1977). The yield/beta/residual risk tradeoff. Working paper 66. Research Program in Finance, Institute of Business and Economic Research, University of California, Berkeley.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 39*, 119–138.

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science, 9*, 277–293.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.

Sharpe, W. F. (1967). A linear programming algorithm for mutual fund portfolio selection. *Management Science, 13*, 449–510.

Sharpe, W. F. (1971). A linear programming approximation for the general portfolio problem. *Journal of Financial and Quantitative Analysis, 6*, 1263–1275.

Stone, B. K. (1973). A linear programming formulation of the general portfolio selection problem. *Journal of Financial and Quantitative Analysis, 8*, 621–636.

Stone, B. K., Adolphson, D. L., & Miller, T. W. (1993). Optimal data selection and grouping: An extension of traditional response surface methodology to observational studies involving non-controlled empirical data generation. *Advances in Mathematical Programming and Financial Planning, 3*, 39–68.

Stone, B. K., & Guerard, J. B., Jr. (2010). Methodologies for isolating and assessing the portfolio performance potential of stock return forecast models with an illustration. In J. B. Guerard Jr. (Ed.), *Chapter 10, Handbook of portfolio construction* (pp. 259–336). New York: Springer.

Stone, B. K., & Guerard, J. B., Jr. (2011). Methodology for testing stock return forecast models for performance and factor alignment/misalignment: Alpha performance potential for Japan. *Proceedings of the 2011 CRSP Forum*. Chicago: University of Chicago.

Stowe, J. D., Robinson, T. R., Pinto, J. E., & McLeavey, D. W. (2007). *Equity asset valuation*. Hoboken NJ: John Wiley & Sons for CFA Institute and Association for Investment Management and Research.

Timmerman, A. (2008a). Elusive return predictability. *International Journal of Forecasting, 24*, 1–18.

Timmerman, A. (2008b). Reply to discussion of elusive return predictability. *International Journal of Forecasting, 24*, 29–30.

Treynor, J. (1965). How to rate the management of investment funds. *Harvard Business Review, 43*(January-February 1965), 63–75.

Treynor, J., & Mazuy, K. (1966). Can mutual funds outperform the market? *Harvard Business Review, 44*(January-February 1966), 131–136.

Williams, J. B. (1938). *The theory of investment value*. Cambridge: Harvard University Press.

Chapter 9
The Duality of Value and Mean Reversion

Noah Beck, Shingo Goto, Jason Hsu, and Vitali Kalesnik

9.1 Introduction

Value investing is the strategy of buying stocks with high fundamentals-to-price ratios and selling stocks with low fundamentals-to-price ratios. A mean reversion strategy, also known as a long-term return reversal strategy,[1] entails buying stocks which have been significantly underperforming and selling stocks which have been outperforming the market over several years. Both value and mean reversion are viewed by market participants as premium-generating strategies. Intuitively, the value and mean reversion strategies are related: Both favor stocks with low prices and dislike stocks with high prices. In the case of value, the low or high prices are appraised relative to company fundamentals. In the case of mean reversion, the prices are evaluated relative to past prices. In this chapter we explore how value and mean reversion are related and what makes them different.

[1]We use the mean reversion terminology to refer to reversal effects throughout this paper.

N. Beck • V. Kalesnik
Research Affiliates, LLC

S. Goto
University of Rhode Island

J. Hsu (✉)
Rayliant Global Advisors, Research Affiliates, LLC, and UCLA Anderson, 620 Newport Center Drive, Suite 900, Newport Beach, CA 92660, USA
e-mail: jason.hsu@rayliantadvisors.com

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_9

We begin with an overview of value and long-term mean reversion in prices, including a comparison of the risk–return characteristics of fundamentals-based and mean reversion investment strategies. (We use the price-to-book-value ratio, or P/B, as the value metric.) Empirical evidence suggests that value and mean reversion benefit from the same source of value added. The performance of the value strategy, however, *dominates* the performance of a mean reverting strategy. We summarize pertinent aspects of Fama and French's (2007) study of stocks' migration across size and style portfolios as an explanation of the small size and value premia. We argue that size, value, and mean reversion strategies outperform as the stock prices migrate from extreme to more average levels. Finally, we explore what it is about the value strategy that gives it an edge over the mean reversion strategy. We demonstrate that, while security prices are mean-reverting, company fundamentals are not. Positive or negative economic shocks are reflected both in company fundamentals and in company price.

By offsetting permanent variations in stock prices driven by corporate fundamentals, the value ratio helps distill the transitory component of the stock price that exhibits mean reversion. For example, we can express the P/B ratio as follows:

$$\begin{aligned}\frac{\text{Price}}{\text{Book Value per Share}} &= \frac{q\times\text{Permanent Component}+\text{Transitory Component}}{\text{Permanent Component}}\\ &= q + \frac{\text{Transitory Component}}{\text{Permanent Component}}\end{aligned}$$

Here q represents the long-run mean of the P/B ratio. Assuming that the level of q has relatively small variation across stocks, the P/B ratio provides a useful signal about the transitory variation of the stock price that is likely to be reversed subsequently.[2] We can also see that the P/B ratio itself exhibits a tendency to revert to its long-run mean level q. Consequently, value ratios like P/B should predict mean reversion (reversal) effects of stock prices.

On the other hand, long-term mean reversion strategies rely on the ratio of a stock's recent price to its past price (or past cumulative returns), a ratio that reflects changes in both the transitory and permanent components of stock prices. This ratio is not as efficient as value ratios in capturing transitory (mean-reverting) variations in stock prices. This is why value ratios contain cleaner information about future price movements than do signals derived from prices alone. The value strategy gains a performance advantage over the mean reversion strategy because it is attuned to a purer signal.

[2]Of course it would be useful to examine the importance of this assumption. For example, if the level of q varies across different industries, then we can use an industry control to improve the return predictability of the P/B ratio. However, because the goal of this paper is to compare general value strategies with general mean reversion (long-term reversal) strategies, we will use a few simplifying assumptions.

9.2 Short-Term Momentum and Long-Term Mean Reversion

It is well known that stock returns are affected by short-term momentum and long-term mean reversion in prices. Short-term momentum can be effectively captured with trading strategies that rebalance portfolio holdings monthly. The basic trading rule is to buy the stocks which outperformed, and sell those which underperformed, over the preceding 12 months.

Researchers investigating long-term mean reversion in stock prices usually build portfolios of the stocks whose prices fell over the 4-year period from 5 years ago to 1 year ago. They skip the most recent year because it is largely dominated by momentum. Long-term mean reversion occurs over a period up to 10 years, but researchers typically choose the shorter timespan over which the effect is strongest. The portfolio would buy the loser stocks, whose prices went down the most in this period, and sell the winner stocks, whose prices concurrently went up the most.[3]

Table 9.1 shows the hypothetical performance statistics for the long-term mean reversion strategy in comparison to other strategies known to have generated significant returns in the past. From January 1931 through August 2013, long-term mean reversion in U.S. equities produced on average about 3.5 % excess return per annum with a Sharpe ratio of 0.35 and a *t*-statistic of 3.14. The risk–reward characteristics of the long-term mean reversion strategy are roughly in line with those of the other major drivers of return.

Practitioners usually do not employ the researchers' simplistic mean-reversion strategy. Instead, they generally look at certain price-to-smoothed-fundamental ratios to capture mean reversion. Academicians have preferred the price-to-book value (P/B) ratio; practitioners have additionally taken into account the ratios of price to the past 5 years of cash flow, sales, and dividends. (Technical analysts

Table 9.1 Risk–return characteristics—U.S. equities (January 1931−August 2013)

	Momentum	Long-term mean reversion	Value	Size	Market
Geometric average excess return (Ann.)	5.9 %	3.5 %	4.5 %	2.9 %	6.3 %
Arithmetic average excess return (Ann.)	7.3 %	4.2 %	5.2 %	3.5 %	7.9 %
Volatility (Ann.)	16.7 %	12.1 %	12.3 %	11.3 %	18.6 %
Sharpe ratio	0.44	0.35	0.42	0.31	0.42
t-Statistic	4.00	3.14	3.80	2.81	3.85

Source: Research Affiliates, LLC, based on data from Kenneth French's website. The value and size factors are HML and SMB, respectively. We use an ad hoc long-term reversal factor for the long-term mean reversion strategy

[3]We borrow the "winner" and "loser" terminology from De Bondt and Thaler (1985).

tend to use the ratio of price to a T-year moving average price.) Table 9.1 shows the results for a value strategy, which buys "value" stocks with high book-to-price (B/P) ratios and sells "growth" stocks with low B/P ratios. The fact that practitioners prefer the value strategy is not surprising. It has a Sharpe ratio of 0.42—appreciably higher than the 0.35 Sharpe ratio registered by the naïve mean reversion strategy.

9.3 Links Between Value and Mean Reversion Strategies

Both the value and the long-term mean-reversion strategies buy stocks with low prices and sell stocks with high prices. The value strategy looks at price-to-fundamental ratios; the mean-reversion strategy looks at current prices relative to past prices. It is not surprising, then, that both strategies should essentially capture the same sources of value-added returns. Table 9.2 displays the correlations between the different return-generating strategies we previously considered. As one would expect, the long-term mean reversion and value strategies have the highest correlation (0.62).

Because value and long-term mean reversion are overlapping strategies, it is natural to ask which is better. We have seen that the value strategy has the higher Sharpe ratio, but that may be due to additional market exposure or an unintentional small stock bias. In order to answer the question directly, we employed a regression-based attribution model whose output is reported in Table 9.3. Panel A shows how the four traditional risk factors contributed to the excess return earned by the long-term mean reversion strategy. The strategy's alpha is not meaningfully different from zero; a significant value factor loading explains the bulk of the mean reversion strategy premium.

Panel B presents a performance attribution analysis from the opposite direction: it traces the value premium to its sources, including the traditional market, size, and momentum factors and an ad hoc long-term mean reversion (LTMR) factor. As anticipated, the value strategy has a significant loading on long-term mean reversion. More importantly, however, mean reversion does not explain the value premium completely. The risk-adjusted alpha is about 4.47 % per annum and is

Table 9.2 Return correlation matrix—U.S. equities (January 1931–August 2013)

	Momentum	Long-term mean reversion	Value	Size	Market
Momentum	1.00				
Long-term mean reversion	−0.23	1.00			
Value	−0.40	0.62	1.00		
Size	−0.15	0.41	0.11	1.00	
Market	−0.34	0.26	0.24	0.34	1.00

Source: Research Affiliates, LLC, based on data from Kenneth French's website

Table 9.3 Factor-based return attribution analysis—U.S. equities (January 1931—August 2013)

Panel A. Long-term mean reversion strategy					
	Alpha (Ann.)	Market	Size	Value	Momentum
Coefficient	−0.69 %	0.01	0.38	0.59	0.05
t-Statistic	−0.71	0.85	14.81	24.66	2.64
Panel B. Value strategy					
	Alpha (Ann.)	Market	Size	LTMR	Momentum
Coefficient	4.47 %	0.04	−0.23	0.64	−0.20
t-Statistic	4.43	2.06	−8.15	24.66	−10.73

Source: Research Affiliates, LLC, based upon data from Kenneth French's website

strongly statistically significant. *When a value strategy anchors on fundamentals, it benefits substantially from additional information that is not captured by long-term mean reversion in prices.*[4]

9.3.1 The Value Premium

Value investing generates a significant premium over cap-weighting as a strategy. Although the interpretation of the premium is vigorously debated, the mechanics are well accepted. Value stocks outperform by becoming less value-oriented over time. That is to say, much of the observed outperformance of the value strategy is driven by the stocks with low price-to-fundamental ratios becoming stocks with higher ones. Value stocks with perennially low price-to-fundamental ratios do not appear to generate meaningful excess returns over time.[5]

More interestingly, value stocks typically migrate toward being "less value"—that is, having a higher price-to-fundamental ratio—not because the denominator falls but because the numerator rises. This observation that value stocks deliver greater performance because they tend to experience price appreciation is the basis for inferring that value characteristics forecast positive price mean reversion. On the other hand, growth stocks' underperformance is substantially driven by their initially high price-to-fundamental ratios drifting lower; the corresponding inference is that growth characteristics forecast negative price mean reversion. Both hypotheses are borne out by empirical observation.

To illustrate the effect of changing P/B value ratios, we reproduce findings from the Fama and French (2007) study, "Migration." Table 9.4 displays the average

[4]These results are consistent with Fama and French (1996) and Daniel and Titman (2006), who show that value effects subsume long-term return reversal effects.

[5]However, Fama and French (2007) state, "value stocks that do not migrate have higher average returns than growth stocks that do not migrate." See Table 9.4.

Table 9.4 Migration effect (June 1927–June 2006)

Portfolio	Average excess return	Average contribution to portfolio's average excess return			
		Minus	Same	Plus	Change in size
Small growth	2.2	−5.3	−1.5	0.5	8.5
Small neutral	5.6	−2.7	0.6	2.6	5.1
Small value	9.2	−0.2	−0.5	4.2	5.6
Big growth	−0.9	−1.2	0.6	0.1	−0.4
Big neutral	1.2	−0.9	0.3	2.2	−0.4
Big value	4.8	0.0	2.3	3.3	−0.7

Source: Fama and French (2007), Table 9.3

excess returns of value, neutral, and growth portfolios composed of small-cap or big-cap stocks. The average returns in excess of the market return demonstrate that, over the 80-year period from June 1927 through June 2006, value stocks outperformed growth stocks and small stocks outperformed big stocks. In addition, each year Fama and French constructed six portfolios holding stocks with similar size and P/B value characteristics (e.g., a small value portfolio) and broke out the returns of portfolios that:

- Moved from growth (high P/B) toward value (low P/B)—the Minus transition
- Stayed in the same size/price-to-book value group
- Moved from value toward growth—the Plus transition
- Moved from small to big or vice versa—the Change in Size transition

Also shown in Table 9.4 are the sources and amounts of positive and negative contributions to the excess return. Small stocks outperform big principally because small stocks migrate to large stock groups as their prices appreciate. The resulting size premium is partially offset by big stocks moving in the other direction. Value stocks outperform growth stocks mostly because Plus transitions due to price increases occur more often for value than for growth stocks, while Minus transitions are more likely for growth stocks. In addition, value stocks that remain in the same portfolio from year to year have higher average excess returns than growth stocks that do not move from one portfolio to another.

Value and size strategies select stocks with extreme prices. The migration study shows that, as prices tend to revert to more average levels, the value and size strategies outperform. In this sense these strategies are related to the mean-reversion strategy, which explicitly bets that extreme prices will tend to move to more average levels.

9.3.2 Using Price Ratios to Predict Mean Reversion Effects

It is instructive to see how using a price-to-fundamental ratio can be an effective approach to forecasting the price mean-reversion effect. Generally, the denominator

Table 9.5 Correlation between book value changes and U.S. stock returns (1962–2012)

	Average cross-sectional correlation between % change in book value and stock return[a]
Correlation	0.11
t-Statistic	8.89

Source: Research Affiliates, LLC, based on data from Compustat and CRSP
[a]1962—2012 (Annual frequency)

of a price-to-fundamental ratio varies slowly and smoothly. In fact, the choice of the fundamental variable that serves as the denominator matters very little, as long as it is significantly less volatile than price and is co-integrated with price over time. Examples of variables that are co-integrated with price, or that trend over time with price, are firm characteristics such as book value or cash flow. Generally, variables that provide information on firm scale satisfy this criterion.

Because stock market participants often exhibit documented behavioral biases, there is a chance they will move prices from the levels that would be purely reflective of companies' fundamental valuations. And there is strong evidence that they do so.

- Shiller (1981) showed that market prices are too volatile to be set by rational investors incorporating information about company fundamentals. His evidence indicated that a significant fraction of stock price changes is due to transitory variations in investors' expectations (fundamentals, discount rates) and hence stock prices contain transitory, mean-reverting, components. Empirical studies by Fama and French (1988) and Poterba and Summers (1988) show that stock prices indeed contain transitory components.
- Roll (1984) studied the U.S. orange juice futures market. Because the demand for orange juice is stable and production is geographically limited, he argued, news about the weather should be the primary driver of price fluctuations. Surprisingly, however, he found that weather related news accounts for very little of the price fluctuations.
- Roll (1988) extended his research to stocks. Using monthly data, he found that at most 35 % of total stock price variation can be explained by publicly available news or by the movement of a close substitute. (Using daily data, that figure dropped to 2 % at most.)

Studies like Shiller's and Roll's confirm that stock price movements come from market participants' aggregate sales or purchases based significantly on considerations other than fundamentals. In other words, many market participants trade on "noise" rather than information, providing depth to the market that facilitates efficient price formation. Meanwhile, their trades may produce transitory deviations in stock prices from their fundamental values.

Price changes have two components: one that is co-integrated with changes in fundamentals, and a mean-reverting component. Table 9.5 shows the cross-sectional correlation between annual U.S. stock returns and changes in book value over the 1962–2012 period. The average correlation, 0.1, is strongly statistically significant.

Table 9.6 Differences in growth rates between U.S. book value winners and losers (1962−2012)

	Fastest-growing companies' subsequent growth rates—slowest-growing companies' subsequent growth rates		
	1 Year	2–5 Years	6–10 Years
Difference in growth	7.94 %	−2.52 %	1.92 %
t-Statistic	3.54	−0.76	0.57

Source: Research Affiliates based upon data from Compustat and CRSP

We have already observed that stock prices are mean reverting. Are fundamental values mean reverting, too?

To answer this question, for each year over the period 1962–2012 we formed portfolios of stocks on the basis of prior growth in book value, and observed the growth in book value over the subsequent year. The portfolios were created using three different versions of prior growth in book value: growth over the previous 1 year, growth from 2 to 5 years ago, and growth from 6 to 10 years ago. Table 9.6 shows the differences in book value growth between "winners" and "losers" from those time periods. (In this context, winners/losers are in the top/bottom 30 % using NYSE breakpoints of book value growth. Specifically, winners are stocks with a greater prior book value growth than the 70th percentile stock on the NYSE. Losers are stocks with a lower book value growth than the 30th percentile stock on the NYSE.)

In the first year after portfolio formation, book values tend to exhibit momentum, but book value growth information older than 1 year has no statistical impact on subsequent growth.[6] This is in contrast to prices, which continue to revert toward the mean for up to 10 years. *Unlike prices, fundamentals do not revert toward the mean.* This result logically implies that fundamentals provide a stable anchor for value strategies.

When a slow-moving co-integrated variable is used in the denominator of the price ratio, most of the change in the ratio over time is driven by price movements. When the price of a stock appreciates significantly, the price ratio also rises, and vice versa. Thus, the change in the price ratio directly measures the past price movement standardized for the company's fundamental scale. The price ratio is the highest for stocks that have had the best cumulative returns relative to their fundamental trend line growth.

The price-to-fundamentals ratio has the additional benefit of encapsulating mean-reversion periodicity; that is, when comparing price ratios among stocks, we need

[6]This cross-sectional evidence is reminiscent of Cochrane's (2005, 2008) time-series evidence that dividend growth is difficult to predict. The result is also related to Daniel and Titman's (2006) finding that "intangible" returns, that cannot be explained by changes in fundamentals, are subsequently reversed.

not be concerned about differences in their mean-reversion periodicity. Moreover, the mean-reversion effect is driven less by periodicity (as seen in momentum strategies) than by the magnitude of the deviation from the long-term trend. If one stock's price ratio is significantly higher (lower) than others', then it is a strong predictor of low (high) future returns because the mean-reversion effect is likely to be stronger for that stock than for other stocks. Thus, selling high P/B stocks, and buying low P/B stocks, is a sound rule for rebalancing a value portfolio. Our research indicates that *a fundamentals-based value strategy stands to gain significantly from long-term mean reversion in prices.*

9.4 Conclusion

Our major findings are easily summarized: The hypothetical long-term mean-reversion strategy results in risk–reward characteristics that are tolerably close to those of established factor-based strategies, and its return series is most highly correlated with the returns of a hypothetical fundamentals-based value strategy. Attribution analysis reveals that the value strategy premium is largely due to a significant mean reversion factor loading; correlatively, the excess return to the mean-reversion strategy proves to have a large value factor loading. However, value strategy performance is superior to that of the mean-reversion strategy. This implies that a value strategy that selects stocks on the basis of fundamentals (in this case, the P/B ratio) reflects more information than a naïve long-term mean-reversion strategy. The evidence presented in Fama and French (2007) demonstrates that the size and value premia are attributable to stocks' transitioning across size/style portfolios. In addition to its inherent interest, this outcome reveals links between the value, size, and mean reversion strategies. Extreme prices tend to converge to more average levels, generating the premia.[7] Finally, we find that company fundamentals themselves are not mean-reverting variables, and that the P/B ratio can be useful in predicting stock returns. We conclude that selling high P/B stocks and buying low P/B stocks is a sound trading guideline—and that a fundamentals-based value strategy can be expected to benefit from long-term mean reversion in stock prices.

Acknowledgement We would like to thank Philip Lawton for his extensive comments and his tireless editorial work.

[7] We have put forth our thesis that the value premium is driven by mispricing in Arnott and Hsu (2008), Arnott et al. (2011), and Chaves et al. (2013).

References

Arnott, Robert, & Jason Hsu (2008). Noise, CAPM and the size and value effects. Journal of Investment Management, 6(1 First Quarter), 1–11.

Arnott, R., Hsu, J., Liu, J., & Markowitz, H. (2011). Can noise create the value and size effects? Research Affiliates Working Paper.

Chaves, Denis B., Hsu, J., Kalesnik, V., & Shim, Y. (2013). What drives the value premium? Risk versus mispricing: Evidence from international markets. *Journal of Investment Management*, *11*(4 Fourth Quarter), 1–18.

Cochrane, J. H. (2005). *Asset pricing* (2nd ed.). New Jersey: Princeton University Press.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies, 21*(4), 1533–1575.

Daniel, K., & Titman, S. (2006). Market reactions to tangible and intangible information. *Journal of Finance, 61*(4), 1605–1643.

De Bondt, W., & Thaler, R. (1985). Does the stock market overreact? *Journal of Finance, 40*(3), 793–805.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy, 96*(2), 246–273.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*(1), 55–84.

Fama, E. F., & French, K. R. (2007). Migration. *Financial Analysts Journal, 63*(3), 48–58.

Poterba, J. H., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics, 22*(1), 27–59.

Roll, R. (1984). Orange juice and weather. *American Economic Review, 74*(5), 861–880.

Roll, R. (1988). R-squared. *Journal of Finance, 43*(2), 541–566.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review, 71*(3), 421–436.

Chapter 10
Performance of Earnings Yield and Momentum Factors in US and International Equity Markets

Jose Menchero and Zoltán Nagy

10.1 Introduction

At an intuitive level, savvy investors have long appreciated the notion of diversification. Not until Markowitz (1952), however, was the concept of trading-off risk and return placed on a firm theoretical foundation. The pioneering work of Markowitz introduced mean-variance optimization as an investment tool and gave birth to Modern Portfolio Theory. The required inputs to solve the Markowitz optimization problem were the expected returns of each asset, the asset covariance matrix, and a specification of investment constraints.

Markowitz defined an "efficient" portfolio as one with the highest expected return for a given level of risk—or equivalently, the lowest risk for a fixed expected return. The set of all such portfolios maps out the so-called efficient frontier. According to Markowitz, rational investors will choose to hold a specific portfolio on the efficient frontier consistent with their risk tolerance.

Tobin (1958) extended the Markowitz theory in a simple yet profound way. By merely treating cash as another investable asset, Tobin showed that all efficient portfolios fell onto a straight line known as the "Capital Market Line." This line connected the risk/return of cash with the risk/return of a "special" portfolio on the efficient frontier holding only risky assets. This special portfolio had the highest expected return per unit of risk of all portfolios on the efficient frontier. This implied that a mean-variance investor would select a portfolio that was a combination of cash and the special portfolio of risky assets on the efficient frontier. This result became known as the two-fund separation theorem.

J. Menchero (✉)
Bloomberg, Pier 3, San Francisco, CA 94111, USA
e-mail: jmenchero@bloomberg.net

Z. Nagy
MSCI, Kassák Lajos utca 19-25, Budapest 1134, Hungary

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_10

Sharpe (1964) extended the ideas of Markowitz and Tobin with the development of the Capital Asset Pricing Model (CAPM). By employing certain assumptions, such as each investor having mean-variance preferences and homogeneous expectations, Sharpe was able to show that the special portfolio on the efficient frontier with the maximum expected return per unit of risk was none other than the market portfolio itself. Furthermore, the CAPM made the important prediction that the expected return of any asset was given by the market beta of the asset multiplied by the expected return of the market portfolio.

Interestingly, in an unpublished manuscript from 1962, Treynor (1962) also arrived at the core result of the CAPM. More specifically, Treynor showed that under the standard assumptions, the expected return of any asset was proportional to its covariance with the market portfolio held collectively by all investors. In other words, high stock volatility did not necessarily imply high expected returns; correlations were also important.

While the CAPM represents an elegant theoretical framework, it has not held up well to empirical scrutiny. Academics began to discover and document several "pricing anomalies" in apparent violation of the CAPM. For instance, Basu (1977) documented the value premium, in which stocks with high earnings yield (E/P ratios) tended to outperform. Similarly, Jegadeesh and Titman (1993) documented the momentum effect, whereby recent "winners" tended to outperform recent "losers." One of the most serious empirical criticisms of CAPM came from Fama and French (1992), who showed that market beta did not appear to have a return premium, in direct violation of the most basic prediction of CAPM.

Many quantitative investors have sought to outperform the market by exploiting such pricing anomalies. Two of the most common "quant signals" are earnings yield and momentum. In this article, we study the performance of these two strategies in both the US and International markets. We consider both pure factor portfolios and optimized factor portfolios. We analyze the period from January 1997 through August 2014. This period contains several interesting events, such as the Internet Bubble, the Quant Meltdown, and the Financial Crisis.

10.2 Pure Factor Portfolios

Pure factor portfolios are very useful for studying the performance of investment strategies, as they isolate the effect of the underlying signal. As described by Menchero (2010), pure factor portfolios are formed by multivariate cross-sectional regression of stock returns against a set of factor exposures. For an individual country, the factors typically consist of styles and industries. Pure style factor portfolios have unit exposure to the particular style, and zero exposure to all other styles and industries. Hence, these portfolios capture the performance of the style in that particular country, net of other styles and industries. In a global setting, country factors are typically added as explanatory variables. In this case, the pure style factor portfolios are neutral to all other styles, industries, and countries.

Pure factor portfolios are not unique as their holdings depend on the regression-weighting scheme. In this paper, we consider three distinct regression-weighting schemes: (a) market capitalization, (b) square-root of market capitalization, and (c) equal. In the cap-weighted case, the pure factor portfolios are naturally dominated by large-cap stocks. By contrast, small-cap stocks dominate the pure factor portfolios for equal-weighted regression. Root-cap weighting is intermediate, assigning more weight to large-cap stocks, but not so much as in the cap-weighted scheme.

Two additional items are required to fully specify the pure factor portfolios. The first is the factor exposure matrix. For US portfolios, we take the exposures from the Barra US Equity Model (USE4). For International portfolios, exposures are taken from the Barra Global Equity Model (GEM3).

The second item we must specify is the estimation universe over which the cross-sectional regressions are to be performed. Our estimation universe is derived from the MSCI All Country World Investable Market Index (ACWI IMI), a broad global equity index spanning both developed and emerging markets. For US portfolios, we carve out the US portion of the index (USA IMI). For International portfolios, we exclude all US stocks from ACWI IMI. All pure factor portfolios are rebalanced on a monthly basis.

For official reporting purposes it is of course essential to use geometrically compounded returns. However, for purposes of understanding the behavior of a portfolio over an extended time period, it is often advantageous to use arithmetic returns. This makes it much easier to compare the performance of the factor over different time periods without the visual distortions that accompany geometric compounding. Therefore, in this paper, we ignore the effects of geometric compounding and simply plot the cumulative arithmetic returns of the portfolios. The annualized volatilities of the portfolios are computed as the standard deviation of monthly returns scaled by the square root of 12. Similarly, the annualized returns are computed as the mean monthly return multiplied by 12.

In Fig. 10.1, we present the cumulative performance of pure earnings yield factor portfolios for the US market. On the whole, the strategies performed well, with cumulative returns over the sample period ranging from 30 to 40 %. The most prominent feature occurred during the Internet Bubble, when earnings yield strongly underperformed prior to the peak (1999) before recovering spectacularly following its collapse (2000–2001). Another notable feature is the poor performance of earnings yield from the time of the Quant Meltdown in August 2007 through the end of 2008. In 2009, earnings yield strongly recovered but performance has been modest since that time. The pure earnings yield strategy returns in U.S. stocks are largely independent of the capitalization-weighting of the variable; see how (a), (b), and (c) lines move in the very similar magnitudes.

While the cumulative performance over the entire sample period was largely similar for the three portfolios, during individual subperiods the differences were sometimes quite significant. For example, during the Internet Bubble the cap-weighted strategy performed much worse than the equal-weighted strategy prior to the peak of the bubble, but recovered much more strongly after its collapse in

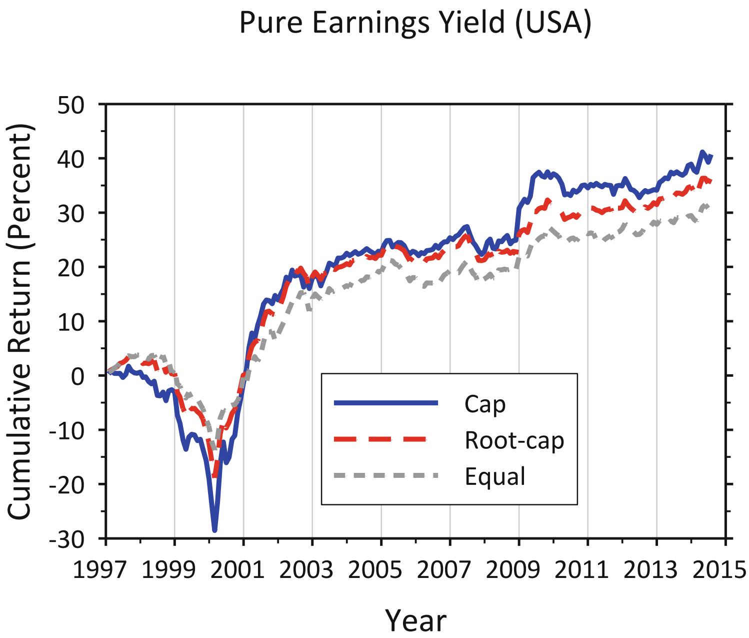


Fig. 10.1 USA pure earnings yield factor portfolios. Cumulative performance is reported for (**a**) market-capitalization regression weights, (**b**) root-cap regression weights, and (**c**) equal regression weights

early 2000. This suggests that during the Internet Bubble, earnings yield was driven primarily by large-cap stocks.

In Fig. 10.2, we plot the cumulative performance of pure earnings yield factor portfolios for International equities. The three strategies tracked each other quite closely over the sample period, suggesting that earnings yield behaved similarly in the large-cap and small-cap segments. By the end of the sample period, cumulative returns in each case exceeded 40 %. What is particularly striking is that International earnings yield did not suffer the same severe drawdown witnessed in the US market during the Internet Bubble. In fact, from the start of the sample period until the peak of the Internet Bubble, all three strategies earned positive cumulative returns, in stark contrast to the performance in the US market. It is also worth noting that International earnings yield performed poorly from the Quant Meltdown through the end of 2008, similar to what was observed in the US market. Since early 2009, International earnings yield has performed quite strongly.

In Fig. 10.3 we plot the cumulative performance of momentum pure factor portfolios in the US equity market. We see that the equal-weighted strategy outperformed the cap-weighted strategy. Furthermore, comparing Figs. 10.3 to 10.1, it is apparent that momentum was considerably more volatile than earnings yield in the US market. Momentum in the US market was characterized by some periods

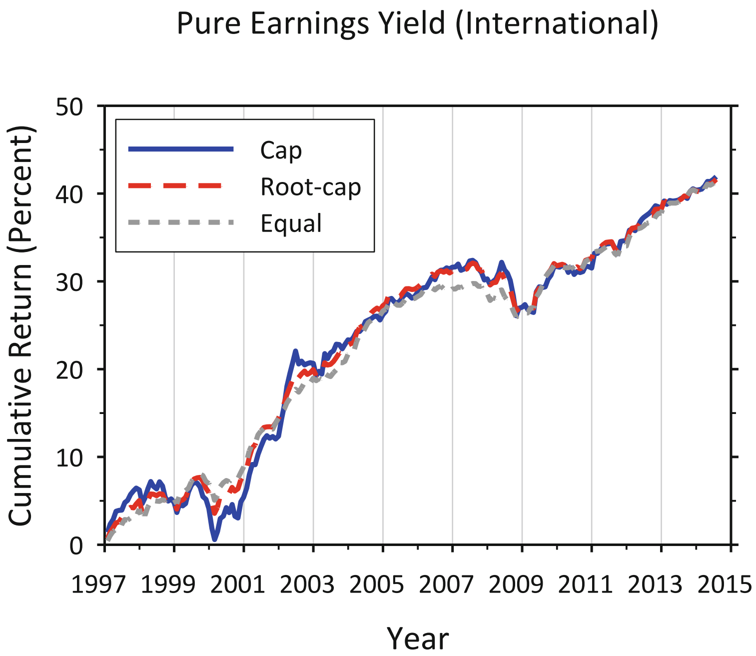


Fig. 10.2 International pure earnings yield factor portfolios. Cumulative performance is reported for (**a**) market-capitalization regression weights, (**b**) root-cap regression weights, and (**c**) equal regression weights

of extremely strong performance (e.g., 1997–2001), and other periods of severe drawdown. The two strongest drawdowns occurred in late 2002 and early 2009, both coinciding with the bottom of bear markets. While US momentum strategies partially recovered during the period 2010–2013, subsequent performance has been flat. Moreover, momentum strategies in the US are still well below their high-water marks of 2008.

It is interesting to consider whether momentum may be prone to crashes near the bottom of bear markets. Such periods are characterized by pervasive pessimism among investors. Most stocks will have negative performance during a bear market, but negative momentum stocks by definition will have performed even worse than average. These are likely to be companies whose prospects appear particularly bleak during the bear market. When the market eventually rebounds, optimism returns and investors may conclude that the negative momentum stocks have been oversold. These stocks are then likely to rebound more strongly than the overall market. Since negative momentum stocks are short in the pure factor portfolio, they can have the effect of "torpedoing" the momentum factor performance.

Daniel and Moskowitz (2013) investigated the performance of momentum strategies across multiple asset classes going back to the Great Depression. They found evidence that momentum strategies were prone to crashing during "panic"

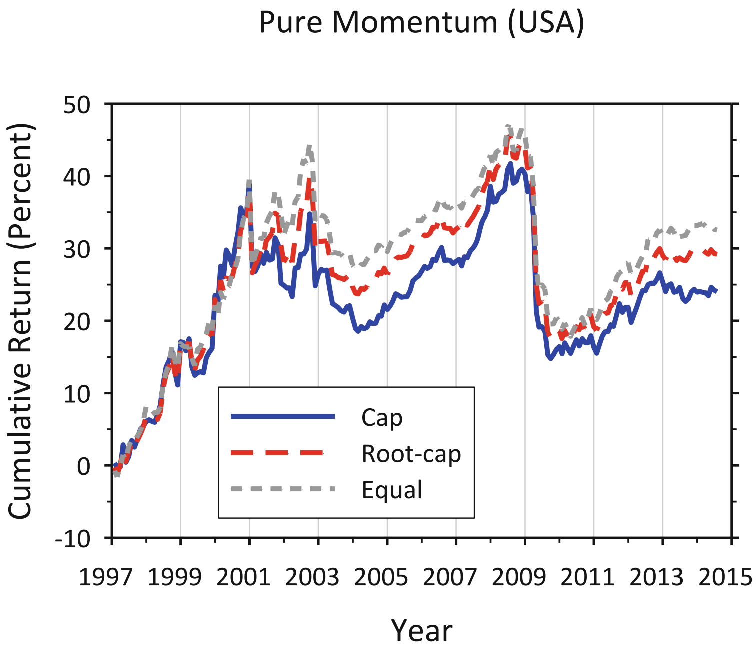


Fig. 10.3 USA pure momentum factor portfolios. Cumulative performance is reported for (**a**) market-capitalization regression weights, (**b**) root-cap regression weights, and (**c**) equal regression weights

states, which correspond to periods following steep market declines coupled with high volatility. They also found that the momentum crashes were contemporaneous with market rebounds.

In Fig. 10.4, we show the cumulative performance of pure momentum in International equity markets. Similar to US momentum, the equal-weighted strategy outperformed the cap-weighted strategy. By contrast, the cumulative performance of momentum in International markets was considerably stronger than in the US. It is also interesting to observe that there was no significant drawdown to International momentum in 2002, unlike the US market. Furthermore, the momentum crash of 2009 was less severe than in the US market. Finally, we see that in 2010 momentum rebounded much more strongly in International markets than for the US market. As a result, the cumulative performance of momentum is at or above the pre-peak levels of 2008.

In Table 10.1, we summarize the risk and return characteristics of the various pure factor portfolios. Several points are worth highlighting. For earnings yield, we see that the cap-weighted strategy outperformed the equal-weighted strategy. For momentum, however, the opposite was true. We also see that the equal-weighted regressions produced the lowest volatility portfolios. This was due to excessive concentration in the portfolios constructed using cap-weighted regression. Finally,

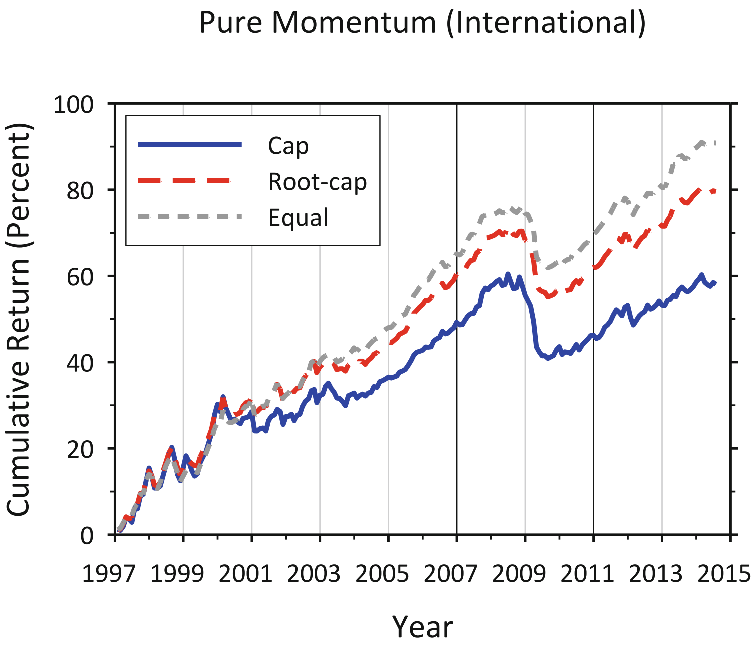


Fig. 10.4 International pure momentum factor portfolios. Cumulative performance is reported for (**a**) market-capitalization regression weights, (**b**) root-cap regression weights, and (**c**) equal regression weights

Table 10.1 Pure factor portfolios under three distinct regression-weighting schemes

Variable	Market segment	Factor name	Cap weight	Root-cap weight	Equal weight
Return	USA	Earnings Yield	2.31	2.08	1.81
Risk	USA	Earnings Yield	5.23	3.72	3.22
IR	USA	Earnings Yield	0.44	0.56	0.56
Return	International	Earnings Yield	2.38	2.37	2.35
Risk	International	Earnings Yield	2.53	1.71	1.48
IR	International	Earnings Yield	0.94	1.38	1.59
Return	USA	Momentum	1.37	1.66	1.85
Risk	USA	Momentum	7.08	6.48	6.39
IR	USA	Momentum	0.19	0.26	0.29
Return	International	Momentum	3.30	4.52	5.16
Risk	International	Momentum	5.05	4.18	3.87
IR	International	Momentum	0.65	1.08	1.33

Factor exposures were equal to 1 in all cases. The sample period is from January 1997 through August 2014 we note that the International strategies had much higher information ratios than their US counterparts. This was due to the combined effect of higher returns and lower risk.

10.3 Optimized Factor Portfolios

Optimized factor portfolios are constructed by mean-variance optimization. The required inputs are (a) the expected returns of the assets, (b) the asset covariance matrix, and (c) a set of investment constraints. The expected returns are often taken as proportional to a factor exposure, or perhaps a combination of factor exposures. The asset covariance matrix is typically obtained by means of a multifactor risk model. For a fixed-exposure constraint, the optimized factor portfolio will have the minimum risk of all portfolios with the same exposure. For a fixed-volatility constraint, the optimized portfolio will have the maximum factor exposure of all portfolios with the same volatility.

Mean-variance optimization employs the risk model to reduce volatility in two ways. First, factor covariances are exploited to hedge the risk of the factor on which the portfolio is tilting. Second, the optimizer uses specific risk forecasts to diversify the idiosyncratic volatility of the portfolio.

We consider optimized portfolios constructed in two distinct ways. The first targets a fixed exposure to the desired factor; the second targets a fixed level of risk. If no further constraints are imposed, then the two portfolios are simply scaled or levered versions of one another. That is, the *ex ante* information ratios of the two portfolios are the same. However, if additional constraints (e.g., turnover) are imposed, then the two portfolios will in general have different *ex ante* information ratios. Even without additional constraints, the ex post information ratios will be different. This is due to the fact that the fixed-risk portfolio must effectively delever during volatile periods to maintain the volatility target, whereas the constant exposure portfolio will experience higher volatility.

We consider several other types of constraints in this paper. The first is the full-investment constraint, which says that the portfolio weights must sum to unity. All portfolios considered in this paper are subject to the full-investment constraint. Consequently, the active portfolios are strictly dollar neutral. The second constraint that we consider is the long-only constraint, which disallows negative weights in the portfolio. The third constraint is the turnover constraint, which limits the amount of trading that can be carried out at each portfolio rebalancing.

Constraints restrict the ability of the optimizer to reduce risk through hedging and diversification. As a result, compared with an unconstrained optimization, the constrained portfolios will have a higher level of *ex ante* risk for a given exposure to the factor. Alternatively, constraints lower the expected return for a given level of risk.

10.4 Unit-Exposure Optimized Portfolios

We now consider the performance of unit-exposure optimized earnings yield portfolios in the US equity market. The earnings yield factor was taken from the Barra US Equity Risk Model (USE4), which was also used to supply the asset covariance matrix. The investment universe was the MSCI USA IMI index.

We consider three distinct optimized factor portfolios. The first portfolio has no constraints other than the full-investment constraint. The second portfolio has the long-only and full-investment constraints, but no limit on portfolio turnover. The third portfolio is also long-only and fully invested, but adds a one-way monthly turnover limit of 4 %.

In Fig. 10.5, we report the cumulative performance of the three optimized portfolios for the US equity market. If we suppose that stock alphas are proportional to earnings yield exposure, then these optimized portfolios all have the same expected return since they have the same exposure to alpha. In addition, they would have the same expected return as the pure factor portfolios presented in Fig. 10.1, which also have unit exposure to the factor. We see that the cumulative performance

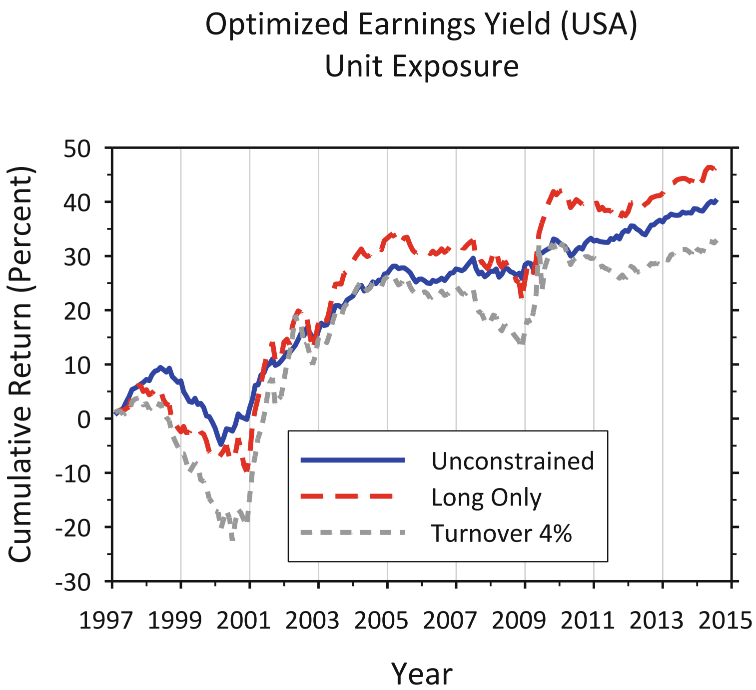


Fig. 10.5 USA optimized earnings yield portfolios (unit exposure). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 4 % monthly one-way turnover

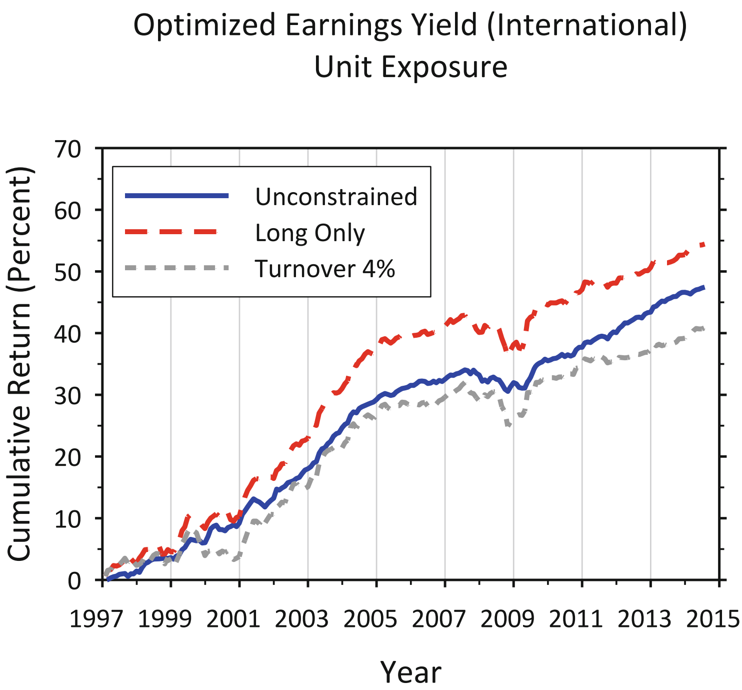


Fig. 10.6 International optimized earnings yield portfolios (unit exposure). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 4 % monthly one-way turnover

of the portfolios in Fig. 10.1 is indeed quite similar to the cumulative returns in Fig. 10.5. Also note that the return patterns in Fig. 10.5 are broadly similar to those in Fig. 10.1. Namely, the earnings yield strategy underperformed during the Internet Bubble, recovered strongly after its collapse, performed poorly between the Quant Meltdown and the end of 2008, and experienced positive returns since early 2009.

In Fig. 10.6, we report the cumulative performance of unit-exposure optimized earnings yield portfolios in the International market. The earnings yield factor in this case was taken from the Barra Global Equity Model (GEM3), which was also used in the optimization. The investment universe was the MSCI ACWI IMI index, excluding USA. Cumulative performance of the optimized portfolios ranged from 40 to 55 %, similar to the performance of the pure factor portfolios in Fig. 10.2. Also note that the broad features of Fig. 10.2 are clearly reflected in Fig. 10.6.

While the realized returns of the three optimized portfolios were similar, the volatilities were dramatically different. The lowest volatility portfolio was the unconstrained one, whereas the constrained portfolios exhibited significantly higher volatility. This is to be expected, since in the unconstrained case the optimizer is better able to diversify and construct risk-reducing hedges.

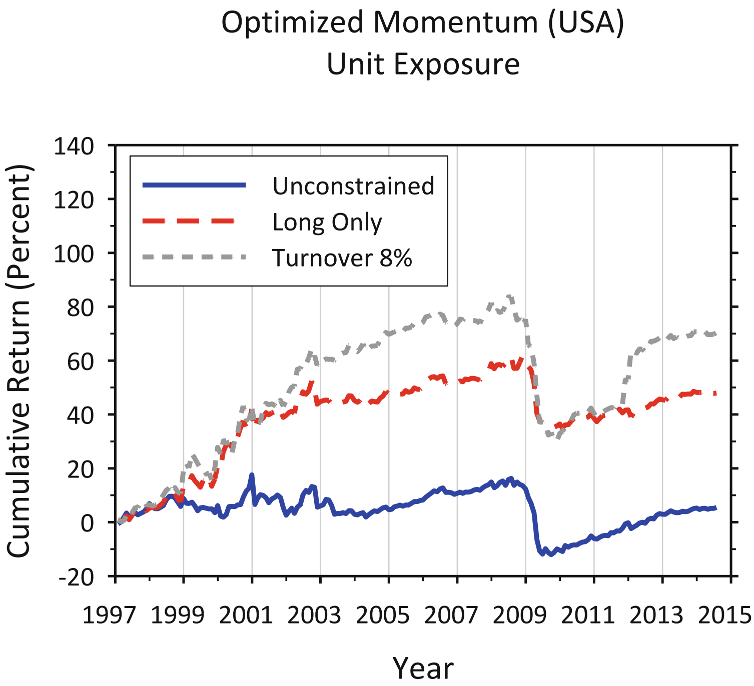


Fig. 10.7 USA optimized momentum portfolios (unit exposure). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 8 % monthly one-way turnover

In Fig. 10.7, we report the cumulative performance of unit-exposure optimized momentum portfolios in the US market. The momentum factor was taken from the Barra USE4 model. In order to better capture the faster-moving momentum signal, we now extend the monthly turnover limit to 8 %. Here we find dramatically different performance for the three optimized portfolios. The unconstrained portfolio had the worst cumulative performance, at about 5 %. By contrast, the portfolio with the long-only and turnover constraints had cumulative returns of roughly 70 %.

There are two possible explanations to account for the large return differences observed in Fig. 10.7. The first possibility is that the expected alphas were not proportional to momentum exposure. In that case, the unconstrained momentum optimal portfolio may inadvertently tilt in a negative alpha direction, thus detracting from the performance.

A second possibility is that this seemingly large difference is attributable to chance. If true alphas are assumed to be proportional to momentum exposure, then the portfolios should have the same expected return. The tracking error between the two portfolios was about 7.7 % annualized. Using square root of time scaling, this

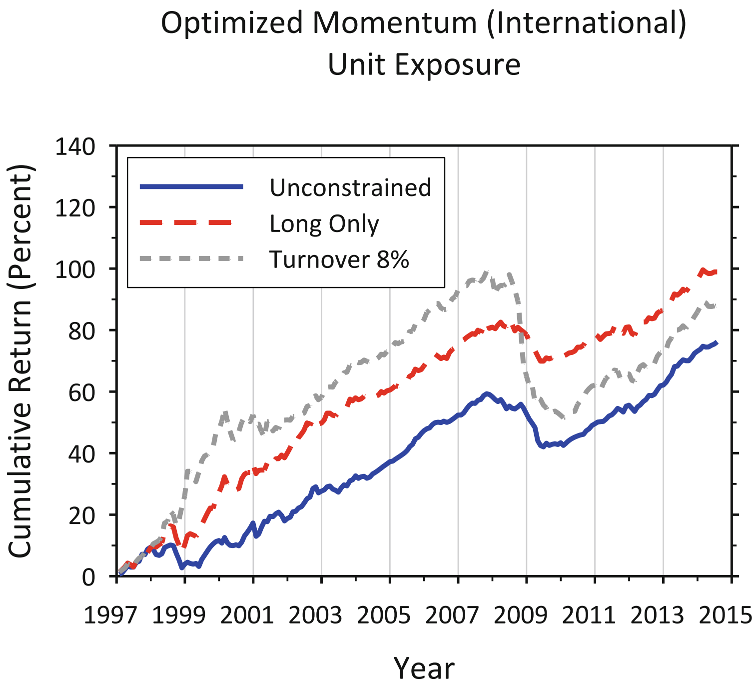


Fig. 10.8 International optimized momentum portfolios (unit exposure). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 8 % monthly one-way turnover

translates into a standard deviation of about 32 % over the roughly 17-year sample period. The observed difference of 65 % represents roughly two standard deviations, thus placing the observation just on the borderline of statistical significance.

In Fig. 10.8, we report the cumulative performance of unit-exposure optimized momentum portfolios in the International market. Here the cumulative returns are similar in magnitude to the returns of the pure factor portfolios in Fig. 10.4. Qualitatively, the return patterns are also broadly similar to those in Fig. 10.4. That is, both figures are characterized by strong performance with a significant drawdown in 2009.

In Table 10.2, we report summary risk and return characteristics for the unit-exposure optimized portfolios. It is interesting to note that in every case, volatility increased as we imposed additional constraints. For example, the US unconstrained earnings yield portfolio had a realized volatility of 2.73 %, versus 4.70% when the long-only constraint was imposed, and 5.94 % when the turnover constraint was applied.

It is also interesting to compare the volatilities of the optimized portfolios in Table 10.2 with the pure factor portfolios of Table 10.1. In every case, the

Table 10.2 Optimized factor portfolios for fixed exposure

Variable	Market segment	Factor name	No constraint	Long-only constraint	LO and TO constraint
Return	USA	Earnings Yield	2.30	2.65	1.87
Risk	USA	Earnings Yield	2.73	4.70	5.94
IR	USA	Earnings Yield	0.84	0.56	0.32
Return	International	Earnings Yield	2.70	3.09	2.32
Risk	International	Earnings Yield	1.38	2.15	2.26
IR	International	Earnings Yield	1.96	1.44	1.03
Return	USA	Momentum	0.34	2.84	4.10
Risk	USA	Momentum	6.10	6.63	8.97
IR	USA	Momentum	0.06	0.43	0.46
Return	International	Momentum	4.45	5.71	4.71
Risk	International	Momentum	3.63	4.31	6.54
IR	International	Momentum	1.22	1.33	0.72

The target factor exposure was equal to 1 in all cases. Monthly turnover constraints were set to 4.0 % for earnings yield, and 8.0 % for momentum. The sample period is from January 1997 through August 2014

unconstrained optimal portfolio had lower volatility than any of the corresponding pure factor portfolios. This indicates that the risk model was successful in reducing portfolio volatility. Of course, once constraints are imposed, there is no guarantee that the risk of an optimized portfolio will be lower than that of a pure factor portfolio. For instance, the 5.94 % volatility for the US earnings yield optimized portfolio with long-only and turnover constraints was greater than the volatilities of the corresponding pure factor portfolios in Table 10.1.

10.5 Fixed-Volatility Optimized Portfolios

We now consider optimized portfolios that are rebalanced each month to have a fixed *ex ante* tracking error. In all cases, we set the tracking error target to 4 %. We again take the factor exposures from the Barra USE4 and GEM3 models. The investment universes are the same as before.

In Fig. 10.9, we report cumulative performance for US optimized earnings yield portfolios. In this case, the unconstrained portfolio will have the largest average exposure to the factor. If alphas are assumed proportional to earnings yield exposure, then the unconstrained optimal portfolio will have the highest expected return. From Fig. 10.9, we see that the realized returns were in fact greatest for the unconstrained optimal portfolio, and least for the most highly constrained portfolio.

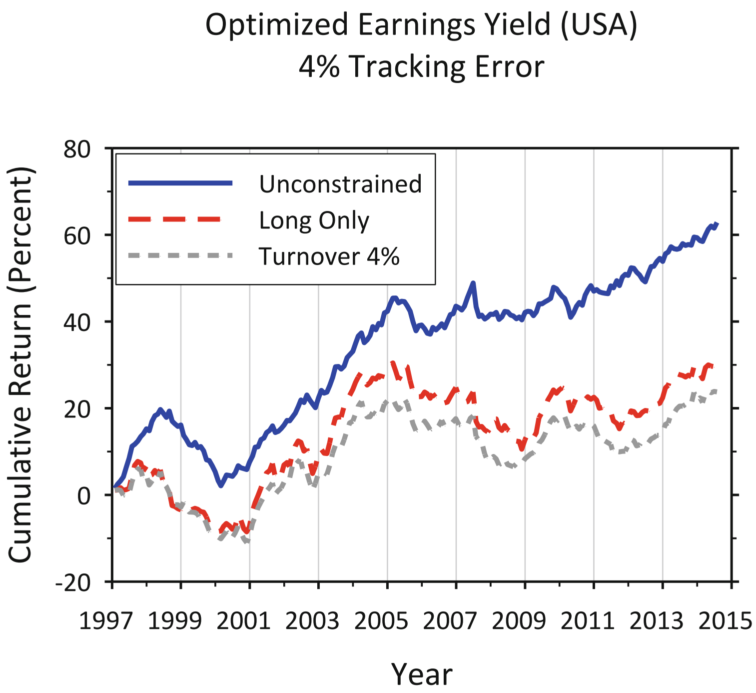


Fig. 10.9 USA optimized earnings yield portfolios (4 % tracking error). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 4 % monthly one-way turnover

In Fig. 10.10, we report cumulative performance for International optimized earnings yield portfolios. Once again, we find that the unconstrained portfolio had the highest realized returns, consistent with expectations. Similarly, the portfolio with both long-only and turnover constraints had the lowest realized returns.

In Fig. 10.11, we report cumulative performance for US optimized momentum portfolios with fixed tracking error of 4 %. In this case, we obtain the unexpected result that the unconstrained portfolio had the worst performance, whereas the most constrained portfolio had the highest realized returns. This is consistent with the results shown in Fig. 10.7.

In Fig. 10.12, we report cumulative returns for International optimized momentum portfolios. Again, the *ex ante* tracking error was fixed at 4 % at the start of each month. Here we find that the unconstrained portfolio had the best performance, but only by a slight margin. It is interesting to compare the cumulative performance of the fixed-risk optimized portfolios in Fig. 10.12 with the fixed-exposure optimized portfolios of Fig. 10.8. It is apparent that the 2009 drawdown was much smaller for the fixed-risk portfolios than for the fixed-exposure portfolios. This is because the fixed-risk portfolio effectively was forced to de-lever when momentum volatility

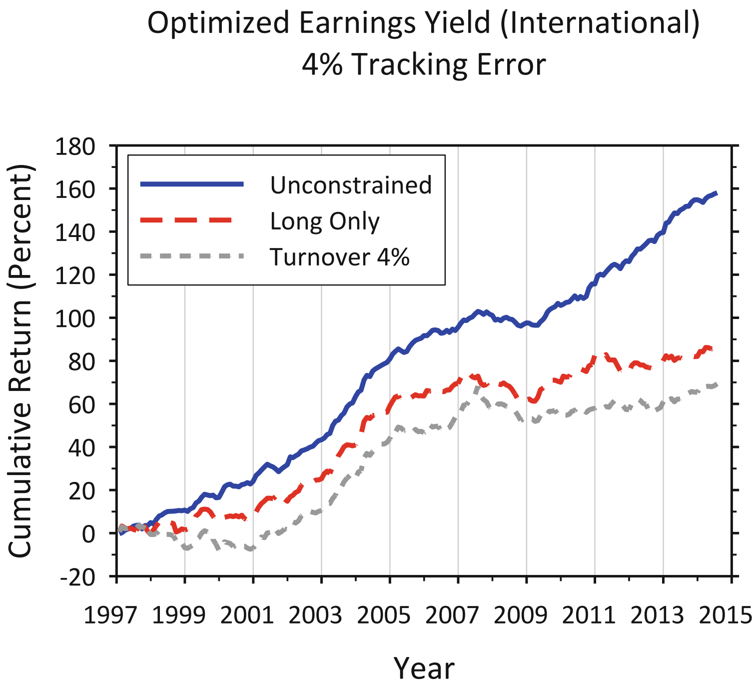


Fig. 10.10 International optimized earnings yield portfolios (4% tracking error). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 4 % monthly one-way turnover

shot up in early 2009. By contrast, the fixed exposure portfolios were forced to "ride out the storm."

In Table 10.3, we report summary risk and return characteristics for the fixed-volatility optimized portfolios. We see that the realized volatilities of the portfolios were in the 4–5% range, close to the predicted risk of 4 %. We also see that the unconstrained portfolios had the highest realized returns, with the exception of US momentum. Finally, comparing the information ratios of Table 10.3 with those of Table 10.2, we find that for momentum they were higher for fixed risk, but for earnings yield they were comparable. We attribute this to the fact that the fixed-risk case for momentum effectively avoided the worst of the drawdown in 2009. Note also that the International strategies had much higher information ratios than their US counterparts.

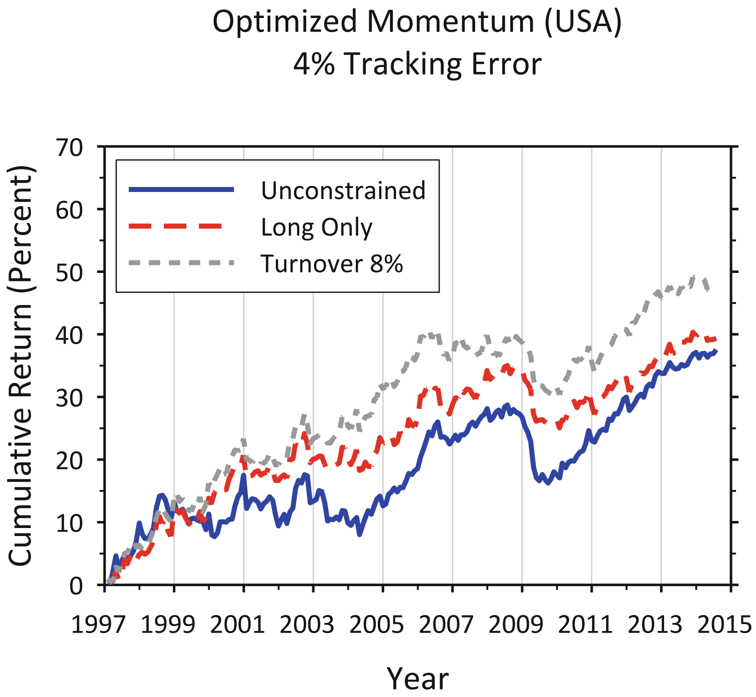


Fig. 10.11 USA optimized momentum portfolios (4 % tracking error). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 8 % monthly one-way turnover

10.6 Summary

We studied the performance of earnings yield and momentum signals in the US and International equity markets. The sample period covered January 1997 through August 2014. We constructed both pure factor portfolios and optimized factor portfolios. For the latter, we considered both fixed-exposure and fixed-volatility portfolios under a variety of investment constraints. In every case, we found that the International strategies outperformed their US counterparts over the sample period. We also found that for the same unit exposure, the unconstrained optimized portfolios had lower realized volatility than the corresponding pure factor portfolios. For the fixed-volatility optimized portfolios, with the exception of US momentum, the unconstrained portfolios had the highest risk-adjusted returns.

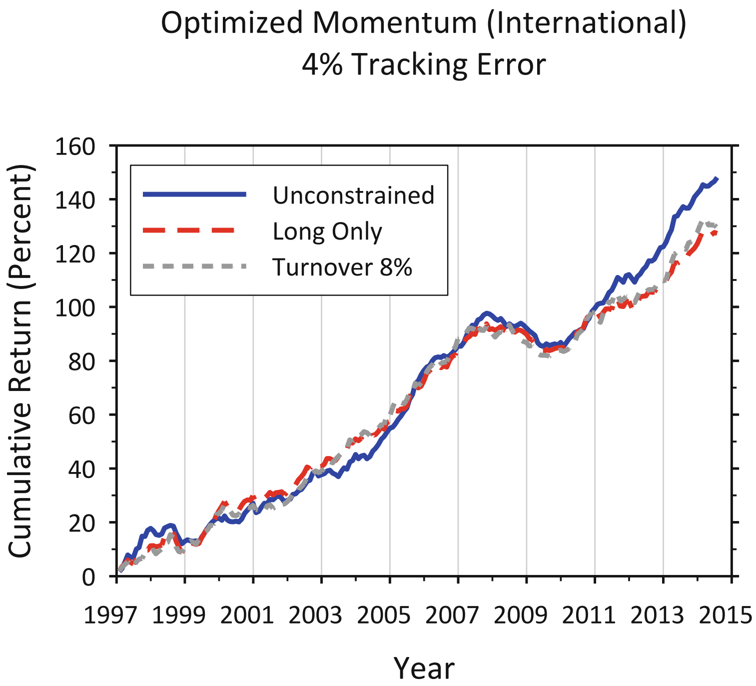


Fig. 10.12 International optimized momentum portfolios (4 % tracking error). Cumulative performance is reported for (**a**) unconstrained, (**b**) long-only constraint, and (**c**) long-only with 8 % monthly one-way turnover

Table 10.3 Optimized factor portfolios for fixed tracking error

Variable	Market segment	Factor name	No constraint	Long-only constraint	LO and TO constraint
Return	USA	Earnings Yield	3.57	1.71	1.35
Risk	USA	Earnings Yield	4.26	5.01	4.37
IR	USA	Earnings Yield	0.84	0.34	0.31
Return	International	Earnings Yield	8.99	4.87	3.94
Risk	International	Earnings Yield	4.03	5.10	5.07
IR	International	Earnings Yield	2.23	0.95	0.78
Return	USA	Momentum	2.14	2.24	2.70
Risk	USA	Momentum	4.41	4.16	4.08
IR	USA	Momentum	0.48	0.54	0.66
Return	International	Momentum	8.42	7.24	7.36
Risk	International	Momentum	4.50	4.66	4.74
IR	International	Momentum	1.87	1.55	1.55

The *ex ante* tracking error target was set to 400 bps annualized for both earnings yield and momentum. Monthly turnover constraints were set to 4.0 % for earnings yield, and 8.0 % for momentum. The sample period is from January 1997 through August 2014

References

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratio: A test of the efficient market hypothesis. *Journal of Finance, 32 (June)*, 663–682.
Daniel, K., & Moskowitz, T. (2013). Momentum crashes. Working paper.
Fama, E., & French, K. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465.
Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance, 48*(1), 65–91.
Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.
Menchero, J. (2010). The characteristics of factor portfolios. *Journal of Performance Measurement, 15*(1 Fall), 52–62.
Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425–442.
Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies, 25*(67), 65–86.
Treynor, J. (1962). Toward a theory of market value of risky assets. Unpublished manuscript.

Chapter 11
Alpha Construction in a Consistent Investment Process

Sebastián Ceria, Kartik Sivaramakrishnan, and Robert A. Stubbs

11.1 Introduction

Markowitz (1952, 1991) developed the mean variance optimization (MVO) model that is widely used in portfolio management. The MVO model achieves an optimal tradeoff between risk and return by solving a quadratic optimization problem that is based on a quadratic utility function that considers the first two moments of asset returns, namely the mean and the variance, to measure the return and the risk of the portfolio, respectively. In addition to trading off risk and return, the MVO model has been extended to include a set of constraints that model additional business requirements imposed by asset owners, regulators, risk managers, and trading desks alike. Additional constraints are also frequently used by portfolio managers to implement investment insights or in order to overcome some of the shortcomings of the MVO model itself.

In this paper we concentrate on studying how these three main ingredients of the MVO model interact with the optimizer to produce optimal portfolios. These inputs to the MVO model include: the alpha vector, representing the expected returns; the risk model, that is used to measure the variance of the portfolio, and a set of constraints. In the traditional quantitative investment process, these three inputs to the MVO model may be developed independently of each other, without much regard to the interaction between them. The main challenge with this *independent* approach to the generation of the optimal portfolio is that this portfolio may not consistently represent the views of the portfolio manager that are expressed in the expected returns. This problem was first addressed more than forty years ago by

S. Ceria
Axioma, Inc., 17 State Street, #2700, New York, NY 10004, USA

K. Sivaramakrishnan • R.A. Stubbs (✉)
Axioma, Inc., 400 Northridge Road, #550, Atlanta, GA 30350, USA
e-mail: rstubbs@axioma.com

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_11

Treynor and Black (1973) where they developed a model that uses "appraisal ratios" to determine the allocation to a market proxy and to each asset. Treynor and Black (1973) said

> If [the investor] has special insights, he will get little, if any, help from the portfolio-balancing literature on how to translate these insights into the expected returns, variance, and covariances the algorithms [of Markowitz] require as inputs.

Since the development of the Treynor-Black model, this problem has been studied by many and evolved considerably (see Black and Litterman 1990, 1992; Grinold and Kahn 2000; Sefton et al. 2004). However, the art and science of constructing relevant inputs for mean-variance optimization remains as important and relevant today as it did when Treynor had such original insights more than forty years ago.

This paper uses the consistent investment process outlined in Stubbs (2013) that builds on earlier work in Black and Litterman (1990, 1992), Grinold and Kahn (2000), and Sefton et al. (2004) to propose a portfolio construction process that takes into account the interaction between all the components of the MVO model to yield a more transparent process that translates superior expected returns into outperforming portfolios. Sivaramakrishnan and Stubbs (2013) showed the advantages of being consistent between the factors used to construct expected returns and those used to construct a risk model, but the consistent investment approach goes beyond that. The consistent investment approach, which is described in more detail in Sect. 11.3, generates the optimal portfolios in three steps: (a) by converting the signals the portfolio manager uses to construct the expected returns into individual portfolios, called factor mimicking portfolios (FMP); (b) by linearly combining the factor mimicking portfolios into a target portfolio, which represents an *idealized* portfolio that optimally combines the signals; and (c) by finding the optimal portfolio as a portfolio that is similar to the target portfolio, but that also satisfies all the additional constraints imposed by the portfolio manager. In contrast to the consistent investment approach, the traditional use of MVO solves a single portfolio optimization problem based on expected returns and risk model that are constructed exogenously to the portfolio construction problem and to each other.

Specifically, in this paper, we illustrate the effectiveness and versatility of the consistent investment approach in enabling the portfolio manager to create a portfolio whose active positions and risks are based on her views and her views only to the extent possible after considering all other frictions such as trading costs and mandated constraints. Moreover, we show how to consider these frictions directly in the three-step consistent process. By considering the constraints that will be in the final portfolio construction problem directly in the computation of our expected returns, we are able to achieve a greater transfer coefficient. Rather than working forecasts of individual asset returns, our examples utilize views based on factors having persistent risk premia. While we do not focus on attribution of returns in this paper, the combination of a greater transfer coefficient and the consistent construction of our MVO inputs leads to greater performance measures such as information ratio, Treynor ratio (see Treynor 1966), and Jensen's alpha (see Jensen 1968). The improvement in these performance measures is a direct consequence of creating an active portfolio that tries to optimally reflect the persistent risk premia of the factors we consider.

The paper is organized as follows: Sect. 11.2 provides a brief overview of the MVO model and discusses various challenges. Section 11.3 describes the main steps of the consistent investment process in detail. The consistent investment process is illustrated on a practical example in Sect. 11.4. Section 11.5 presents some of our conclusions. We discuss some technical details in Appendix A.

11.2 Mean Variance Optimization

There are three main ingredients to an MVO model: the α vector representing the expected returns, the risk model Q, and the constraint set $\mathcal{C}$ that defines the strategy employed by the portfolio manager. Consider a portfolio with an investment universe of n assets. Let h_i denote the weight (proportion of total funds) invested in the ith asset. Let α_i denote the portfolio manager's estimate of the expected return for the ith asset.

We define a *signal* or *factor* as a vector of asset characteristics that explains the cross-section of returns. Throughout this paper, we will differentiate between two types of factors: *alpha* factors that exhibit a long term predictable return trend, and *risk* factors that do not exhibit such a predictable trend. Examples of *alpha* factors include value, momentum, and growth while examples of *risk* factors include industries and countries.

Let X_A and X_R denote the asset exposures to the alpha and risk factors, respectively.[1] Let m be the number of alpha factors and let k be the number of risk factors. Assume that the portfolio manager constructs the overall alpha factor from m different factors in X_A. The risk model is given by

$$Q = X\Omega X^T + \Delta^2 \tag{11.1}$$

where

$$X = \begin{bmatrix} X_A \; X_R \end{bmatrix}$$

is the combined matrix of factor exposures,

$$\Omega = \begin{bmatrix} \Omega_A & \Omega_{AR} \\ \Omega_{RA} & \Omega_R \end{bmatrix}$$

[1]The cross sectional regression model has the form

$$r = X_A f_A + X_R f_R + r_{res}$$

where r, f_A, f_R represent the excess (over the risk free rate) asset, alpha factor, and risk factor, returns, respectively. So, the exposures X_A and X_R actually represent the *betas* or sensitivities of the assets to the alpha and the risk factors in the model.

is the factor covariance matrix, and Δ^2 is a diagonal matrix of specific variances. The overall MVO model is given by

$$\max_{h \in \mathcal{C}} \alpha^T h - \frac{\lambda}{2} h^T Q h, \tag{11.2}$$

where $\lambda > 0$ is an appropriate risk aversion parameter and $\mathcal{C}$ contains the different constraints in the MVO model.

We conclude this section by introducing some concepts that will be used later in the paper. The information ratio (IR) of a portfolio is a measure of the risk-adjusted return of the portfolio, namely, the expected active return of the portfolio divided by the standard deviation of the expected active return. Given a portfolio h^* and a risk model Q, the *implied alpha* is the alpha vector that would yield h^* as the solution to the MVO problem without any constraints. Let $\tilde{\alpha}$ denote the implied alpha of the portfolio. It is easy to demonstrate that

$$\tilde{\alpha} = Qh^*. \tag{11.3}$$

Clarke et al. (2006) developed the Transfer Coefficient (TC), which is used to measure the efficiency with which the alpha signal is transferred to the optimal portfolio. The transfer coefficient (TC) is given by

$$\begin{aligned} \mathrm{TC} &= \frac{\alpha^T h}{\sqrt{\alpha^T Q^{-1} \alpha}\sqrt{h^T Q h}}, \\ &= \mathrm{corr}(Q^{-1/2}\alpha, Q^{1/2}h). \end{aligned} \tag{11.4}$$

The TC can be interpreted as the correlation between the risk-adjusted alpha, $Q^{-1/2}\alpha$, and the risk-weighted portfolio, $Q^{1/2}h$.[2] The TC also represents the correlation between the risk-weighted final portfolio and the risk-weighted unconstrained MVO portfolio (where there are no constraints). In a sense, the TC measures how close the final portfolio is to the unconstrained MVO portfolio. Constraints such as turnover and asset bounds in a realistic strategy lower the TC from its ideal value of 1.

11.3 The Consistent Investment Process

There are three main steps in the consistent investment process proposed in Stubbs (2013). We first provide a brief description of these steps and then discuss them in detail in Sects. 11.3.1, 11.3.2, and 11.3.3.

[2]TC represents the correlation between $Q^{-1/2}\alpha$ and $Q^{1/2}h$ only if the means of α and h are zero.

1. **Transform each alpha signal into factor mimicking portfolios:** Each alpha factor (signal) is transformed into a factor mimicking portfolio (FMP). The FMP is generated by solving an optimization problem which minimizes some measure of the portfolio risk while controlling the exposure of the FMP to other risk and alpha factors. The goal of this optimization problem is to have the FMP replicate the alpha factor returns, while neutralizing other factor exposures. Because we use an optimization model to construct the FMP, we can add constraints to this problem in order to generate a more realistic FMP. We will illustrate the benefits of adding constraints to this optimization problem in Sect. 11.4.
2. **Combine the FMPs into a target portfolio:** The second phase of the consistent investment process is combination of the individual FMPs into a portfolio called the *target portfolio*. We propose to construct this portfolio through the solution of an optimization problem that trades off the risk and expected return of the FMPs. Once again, the procedure is designed to be highly flexible and allows for the inclusion of constraints in order to generate a more realistic target portfolio.
3. **Solve the final portfolio construction problem**: While the target portfolio generated in Step 2 is the portfolio that the PM ideally would like to hold, it may not be investable and/or violate some of the other constraints that are part of the PM's mandates. The goal of this step is to generate a realistic portfolio that is as *similar* as possible to the target portfolio while satisfying all of the additional practical considerations and constraints. We solve this problem by using an MVO model once again with the following characteristics: (a) The expected returns vector is the implied alpha of the target portfolio, (b) The risk model is the matrix Q, and (c) the constraints include the relevant implementation constraints and any additional constraints that are part of the PM's mandate.

11.3.1 Transforming Each Alpha Signal into Factor Mimicking Portfolios

A factor mimicking portfolio (FMP) is a long-short, dollar-neutral portfolio that *represents* a factor. Fama and French (1992) describe simple procedures based on fractile analysis to construct such portfolios, which are also called *Fama-French portfolios* or *FF* portfolios. Although their approach to the generation of FMPs is heuristic in nature, Fama and French (1992) also neutralize other factor exposures when generating FF portfolios. For example, Fama and French (1992) neutralize their size (SMB) factor to the momentum and book/price attributes. Practitioners use similar techniques to neutralize their alpha signals to industry and country factors. For example, Asness (1997) neutralizes value (book to price) to industries by subtracting the market-cap weighted industry book to price average from each assets' (within that industry) book to price value.

Let m be the number of alpha signals considered by the portfolio manager. These alpha signals are represented as columns of the factor exposure matrix X_A in the risk model. The FMP associated with the jth signal is the solution, h^j, to

$$\begin{aligned} \min\ & h^T W h, \\ \text{s.t.}\quad & X_R^T h = 0, \\ & X_A^T h = e_j, \\ & h \in \bar{\mathcal{C}} \end{aligned} \tag{11.5}$$

where W is an appropriate weighting matrix, e_j is the vector with 1 in the jth position and zeros elsewhere, and $\bar{\mathcal{C}}$ contains a set of additional constraints. We will define a *pure FMP* as a dollar-neutral portfolio that has minimum total risk, has unit exposure to the alpha factor, and is neutral (zero exposure) to the other alpha factors as well as all the risk factors. For a pure FMP, the set $\bar{\mathcal{C}}$ is empty.

We can build an FMP that is not *pure* through alternative formulations of (11.5) along the following lines:

1. **Choosing the set of *neutral* factors:** A pure FMP is neutral to all the other factors in the risk model. Alternatively, we can impose that the FMP be neutral to only some of the risk factors in X_R and X_A. We show in the technical appendix that neutralizing for the industry factors in (11.5) is equivalent to some of the common heuristic industry purification schemes used in practice. The following are some alternatives we consider in the solution to (11.5):
   (a) The final portfolio may be required to have null or very small (active) exposure to some of the factors in the risk model. In this case, it may be beneficial to make each FMP neutral to such factors.
   (b) A particular factor may negatively contribute to the return of the final portfolio. In this case, it is better to neutralize the exposure to that factor in the generation of the FMP.
2. **Choosing the weighting matrix W:** We mention three popular choices:
   (a) W is the identity matrix, which orients the FMP towards an equal-weighted quantile spread portfolio.
   (b) $W = M^{-1}$ where M is a diagonal matrix whose entries are the asset market capitalizations. This orients the FMP towards a market-cap weighted quantile spread portfolio, where assets with a larger market capitalization are preferred.
   (c) $W = Q$, where Q is the risk model used in the portfolio construction process. In the case of pure FMPs, the covariance matrix $W = Q$ in (11.5) can be replaced with its diagonal specific variance component $W = \Delta^2$. On the other hand, if we do not neutralize exposures to all other factors, the resulting FMP may be unintuitive because of the correlations in Q. For example, a value FMP constructed in this way could well take large negative exposures to assets with high book to price values.

3. **Incorporating additional constraints:** Constraints from the final portfolio construction problem can also be added to (11.5). We mention two examples below:
   (a) The exposure of the FMP to some additional factors can be controlled if the exposure of the final portfolio to that particular factor may have a negative impact on the return of the final portfolio.
   (b) If signals have predictive abilities which are dependent on the investment horizon, we may require that the different FMPs have exposures which are related to the investment horizon of those signals. For example, Qian et al. (2007) and Gerard et al. (2012) introduce the concept of *horizon IC* to measure the strength and the persistence of the alpha signal.

11.3.2 Combining Factor Mimicking Portfolios into a Target Portfolio

The second phase of the consistent investment process linearly combines the FMPs of the m different alpha signals h^j into the target portfolio. Let $w_j, j = 1, \ldots, m$ be a given set of weights. We call

$$h^{tp} = \sum_{j=1}^{m} w_j h^j \tag{11.6}$$

the target portfolio. The weights are determined by optimally trading off the risk and the return of the different FMPs in an MVO framework.

Given a time series of returns for an FMP for each alpha signal, there are several ways to estimate the return $E[f^i]$ for each FMP and the covariance matrix Θ across all FMPs. We describe our choice below:

1. $E[f^i] = \dfrac{1}{T}\sum_t (r^t)^T(h^{it}),\ i = 1, \ldots, m$, where T is the total number of time periods; and r^t and h^{it} denote the time series of realized asset returns and FMP holdings, respectively.
2. Let Q denote the custom risk model that is used in portfolio construction. We have

$$\Theta_{ij} = (h^i)^T Q (h^j),\ \ i,j = 1, \ldots, m.$$

   Note that

$$(h^{tp})^T Q h^{tp} = w^T \Theta w.$$

The optimization problem that generates the target portfolio can be written as

$$\max_{w} \sum_{i=1}^{m} E[f^i]w_i - \lambda w^T \Theta w \tag{11.7}$$

where $\lambda > 0$ is an appropriate risk threshold. We prefer to use the following equivalent formulation

$$\begin{aligned} \max_{w,h^{tp}} \quad & \sum_{i=1}^{m} E[f^i]w_i - \lambda (h^{tp})^T Q h^{tp}, \\ \text{s.t.} \quad & h^{tp} - \sum_{i=1}^{m} w_i h^i = 0 \end{aligned} \tag{11.8}$$

for the target portfolio as it better highlights the relationship between the optimization problems in the three stages of the consistent investment process. Note that one can also force the target portfolio to satisfy some of the additional constraints in $\mathcal{C}$. If this is the case, then one can also add the constraints $h^{tp} \in \mathcal{C}_{tp}$, where $\mathcal{C}_{tp}$ is subset of the constraints in the final portfolio construction problem, to the MVO optimization problem (11.8).

3. If the FMPs h^i used to generate the target portfolio are all pure, then the FMP returns represent the underlying alpha factor returns if the alpha factors are also in the risk model. These factor returns are usually constructed from a cross-sectional regression model. In this case, one can set $\Theta = \Omega_A$ in (11.7), where Ω_A is the factor covariance matrix constructed from the factor returns. The corresponding problem (11.8) has Q constructed from the factor portion $X\Omega X^T$ of the custom risk model Q.

We must emphasize that for the target portfolio optimization problem (11.8) we only have m unknowns, i.e., the number of FMPs that need to be combined to form the target portfolio. Usually, $m << n$, where n is the number of assets in the portfolio. As a result, *asset-level* constraints in the target portfolio problem may be too restrictive.

11.3.3 Solving the Portfolio Construction Problem

The target portfolio generated in the second phase of the consistent process can violate some of the constraints in the final portfolio construction problem which are included in $\mathcal{C}$. The third phase of the consistent investment process constructs a final active portfolio that satisfies all the constraints in $\mathcal{C}$ and is as close to the target portfolio as possible. This problem is solved with an MVO model that uses the implied alpha of the target portfolio as the vector of expected returns. In other words, the vector of expected returns is chosen as

$$\tilde{\alpha}^{tp} = Qh^{tp} \tag{11.9}$$

where Q is the risk model in 11.1. The portfolio optimization in the third phase can be written as

$$\max_{h \in \mathcal{C}} \; (\tilde{\alpha}^{tp})^T h - \tfrac{\lambda}{2} h^T Q h \tag{11.10}$$

where $\lambda > 0$ is an appropriate risk aversion parameter and $\mathcal{C}$ contains all the constraints in the final portfolio optimization problem. One can easily show (by completing the square) that (11.10) can also be written as

$$\min_{h \in \mathcal{C}} \; \tfrac{1}{2}(h^{tp} - \lambda h)^T Q (h^{tp} - \lambda h). \tag{11.11}$$

This, in turn, implies that choosing the alpha as the implied alpha of the target portfolio actually gives us a final portfolio h that is close to a multiple of the target portfolio h^{tp}.

11.4 Illustrative Example

Consider a portfolio manager who wants to combine three alpha signals in a consistent investment process. The three alpha signals are *Value* (Sales to Price), *Momentum* (assets cumulative return over the last 250 trading days), and *Quality* (Return on Equity). The asset universe and the benchmark is restricted to be the FTSE All-World. The objective is to maximize the expected return subject to

1. Long-Only and Fully Invested.
2. Round-Trip (two-way) Turnover restricted to be at most 15 % per month.
3. Active Predicted Beta Bounds of $\pm 2\,\%$.
4. Active Industry and Country Bounds of $\pm 2\,\%$.
5. Maximum Predicted Active Risk of 3 %.
6. Axioma Style Exposure Bounds of $\pm 10\,\%$ on Liquidity, Leverage, Size, Exchange-Rate Sensitivity, and Volatility.

Our backtest period is from December 2000 to August 2012 where the portfolio is rebalanced at the end of each month. We construct a custom risk model using Axioma's Risk Model Machine (RMM) that also includes the Value, Momentum, and Quality alpha signals.

We now discuss the three stages of the consistent investment process as applied to this example.

1. The Value, Momentum, and Quality FMPs are constructed with $W = M^{-1}$ where M is the diagonal matrix whose entries are the asset market-caps.

   (a) The Value FMP is neutral to countries.

Table 11.1 Factor return contribution in dollar-neutral value FMP

Source of return	Contribution	IR
FMP	2.49 %	0.51
Factor Contribution	3.61 %	0.64
Axioma Style	2.21 %	1.27
Custom Style	2.57 %	0.72
Country	−1.72 %	−0.48
Industry	0.70 %	0.37
Currency	−0.15 %	−0.08
Market	0.00 %	−0.31
Specific Return	−1.12 %	−0.50

(b) The Momentum FMP is neutral to industries.
(c) The Quality FMP is neutral to industries and size.

We will motivate our choices for these FMPs below.

2. The optimization problem that generates the target portfolio is

$$\begin{aligned} \max_{w \geq 0, h^{tp} \in \mathcal{C}_{tp}} \sum_{i=1}^{3} E[f^i] w_i + (h^{tp})^T Q h^{tp}, \\ \sqrt{(h^{tp})^T Q h^{tp}} \leq 3\,\%, \\ h^{tp} - \sum_{i=1}^{3} w_i h^i = 0, \end{aligned} \tag{11.12}$$

where we additionally impose that h^{tp} belong to the set $\mathcal{C}_{tp}$. We will highlight some of our choices of this set later in this section. Note that we impose a risk constraint on the target portfolio that is also one of the constraints in our final strategy. Since we expect each of the alpha signals to add value, we also impose a non-negativity restriction of the FMP weights. The $E[f^i]$ is taken to be the long-term average return for the three FMPs and is constant for the backtest.

3. The expected return for the final portfolio problem is the implied alpha of the target portfolio.

Let us highlight how we arrived at our choice for country neutrality constraints for the Value FMP. For each time period, we first constructed a simple dollar-neutral Value FMP with $W = M^{-1}$ with no factor neutrality constraints. We then ran Axioma's factor Performance Attribution Tool to analyze the resulting FMP. The results are presented in Tables 11.1 and 11.2. Table 11.1 indicates that the FMP has a large negative return betting on countries. Table 11.2 further indicates that the FMP is taking significant positive and negative exposures on countries. For example, the FMP has a large negative exposure to Japan (−14.15 %) and a large positive exposure (8.00 %) to UK. So, we made the FMP neutral to all the countries to force it to bet on stocks within countries rather than on individual countries. The resulting FMP is the country-neutral value FMP and it is the value FMP that

Table 11.2 Factors with the largest exposures in dollar-neutral Value FMP

Source of return	Return contribution	Avg exposure
Notable factors		
Volatility	1.51 %	−8.28 %
Value	2.61 %	100 %
China	−0.08 %	−2.34 %
France	−0.24 %	5.59 %
Germany	−0.20 %	5.81 %
Japan	−0.28 %	−14.15 %
Korea	−0.07 %	−2.16 %
UK	−0.33 %	8.00 %
USA	0.11 %	8.10 %
Banks	0.05 %	−6.74 %
Food & Staples	−0.12 %	4.78 %
Oil, Gas & Consumable Fuels	0.54 %	6.47 %

Table 11.3 Comparing dollar-neutral and country-neutral value FMPs

	Country-neutral value FMP			Dollar-neutral value FMP		
Return	Contribution	Risk	IR	Contribution	Risk	IR
FMP	3.87 %	4.90 %	0.79	2.49 %	4.90 %	0.51
Factor	4.92 %	5.21 %	0.94	3.61 %	5.67 %	0.64
Axioma Style	1.51 %	1.84 %	0.82	2.21 %	1.74 %	1.27
Custom Style	2.57 %	3.61 %	0.71	2.67 %	3.59 %	0.72
Country	0.00 %	0.00 %	—	−1.72 %	3.59 %	−0.48
Industry	0.83 %	2.26 %	0.37	0.70 %	1.91 %	0.37
Currency	0.00 %	0.00 %	—	−0.15 %	1.84 %	−0.08
Market	0.00 %	0.00 %	—	0.00 %	0.00 %	—
Specific Return	−1.04 %	2.06 %	−0.51	−1.12 %	2.26 %	−0.5

we will use in the rest of the section. Table 11.3 compares the dollar-neutral and the country-neutral value FMPs. Note that country-neutral FMPs are also dollar-neutral. Enforcing country-neutrality removes the negative contribution to the return from the country factors. The portfolio IR improves from 0.51 to 0.79. We experimented with different neutrality settings for the Value, Momentum, and Quality FMPs. The IRs for these different settings are summarized in Table 11.4. The Value FMPs perform better with the market-cap weighting. We chose the FMP that was neutral to the country factors; this had the best IR in Table 11.4. For Momentum, we first constructed a dollar-neutral FMP with market-cap weighting. A factor performance attribution analysis indicated that this FMP did not take any sizeable exposures to the industry and country factors. We settled on the Momentum FMP with market-cap weighting that is neutral to the industries since available literature (see Asness (1997)) indicates that Momentum performs better with industry neutrality. Similarly, for Quality, we first constructed a dollar-neutral FMP with market-cap

Table 11.4 IRs for different value, momentum, and quality FMPs

FMP	Weighting (W)	IR			
		None	IndNeutral	CountryNeutral	IndAndCountryNeutral
Value	M^{-1}	0.50	0.20	0.78	0.60
Value	I	0.12	0.05	0.70	0.56
Momentum	M^{-1}	0.30	0.24	0.24	0.12
Momentum	I	0.36	0.26	0.36	0.22
Quality	M^{-1}	0.56	0.53	0.44	0.28
Quality	I	0.77	0.74	0.52	0.48

Table 11.5 Base backtest summary

	Base
Annualized active return	2.27 %
Avg annualized active risk	2.83 %
Avg turnover	14.89 %
Information ratio	0.801

weighting. A factor PA analysis indicated that this FMP had a 10.78 % exposure to the size factor that translated into a negative return of –0.48 %. Moreover, the Quality FMP did not take any sizeable exposures to the industry and country factors either. We finally chose the Quality FMP with market cap weighting that is both industry and size neutral.

We ran our *Base* backtest with these FMPs and following the consistent investment process that we outlined earlier. The results are summarized in Table 11.5. Figure 11.1 compares the average active long holdings of the final portfolio to the average six month horizon ICs for the three alpha signals in each sector of the economy. Note that the quality of the three alpha signals varies substantially in each sector. Also, note that average six month sector ICs for the Value signal are very poor and so we focus on the strength of the Momentum and the Quality signals in each sector. These signals both appear to be very strong in the *Telecommunication Services* sector but the *Base* portfolio seems to take relatively small bets in this sector. In general, we see that the sector bets of the *Base* portfolio are not proportional to the sector ICs of the Momentum and Quality signals. This is not surprising since the active beta, industry, and country bound constraints require the portfolio to take very small exposures to these factors. As a result, the portfolio bets are concentrated in certain sectors or countries. So, we add additional constraints to the final strategy that limits the long holding in each sector to be at most 5 % of the portfolio size. We refer to these new constraints as the *active share* constraints. We ran a second backtest called *Base* + *CAS* (CAS = Constrained Active Share) that is identical to the base strategy with the only difference being the active share constraints in the final strategy. The results of this backtest are given in Table 11.6. Note that the IR of the portfolio has improved from 0.8 to 0.85. This is the typical process employed by most portfolio managers where additional constraints are added in the final portfolio problem to overcome some shortcomings. The addition

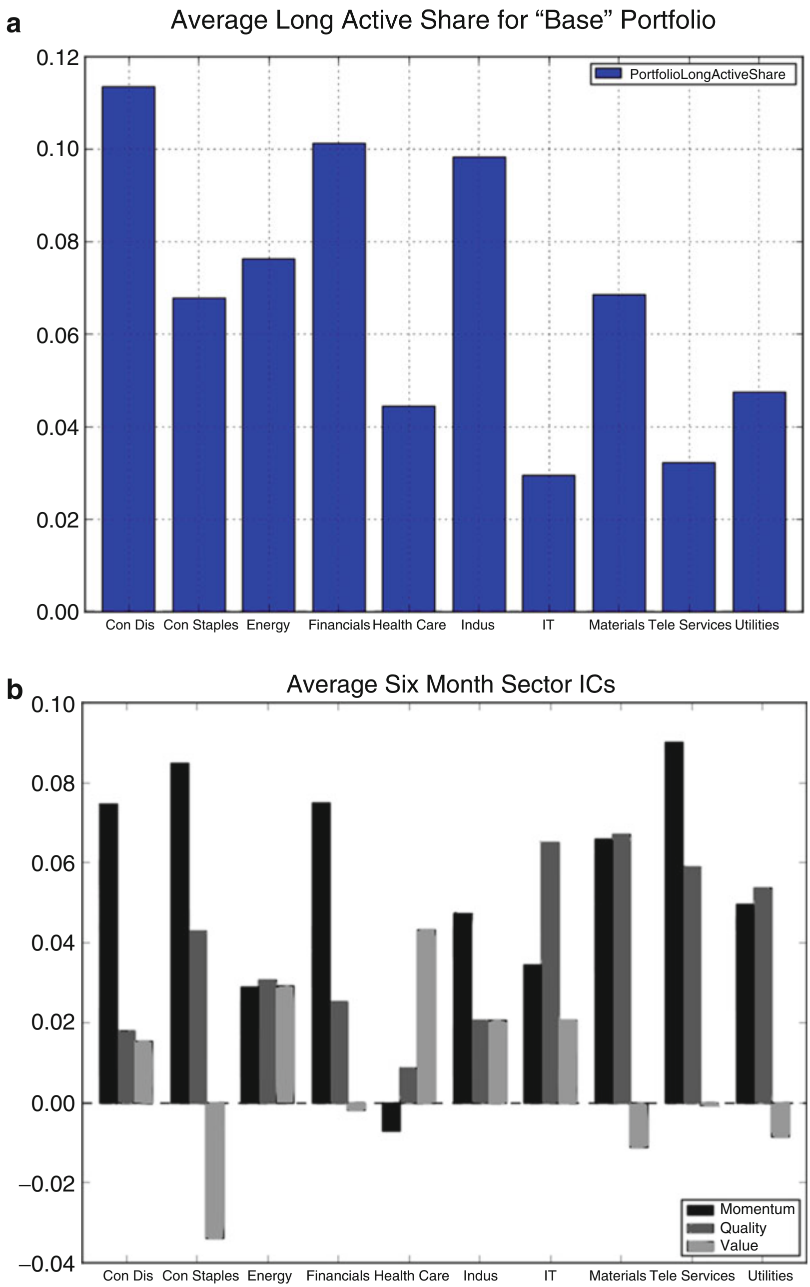


Fig. 11.1 Comparing the average long active holdings to the average six month horizons ICs for alpha signals in each sector. (**a**) Active Share in each sector. (**b**) Average six month sector ICs

Table 11.6 Base + CAS backtest summary

	Base + CAS
Annualized active return	2.36 %
Avg annualized active risk	2.78 %
Avg turnover	14.89 %
Information ratio	0.85

Table 11.7 BaseCFMP backtest summary

	BaseCFMP
Annualized active return	2.95 %
Avg annualized active risk	2.94 %
Avg turnover	14.89 %
Information ratio	1.01

of the constraints in the final portfolio, however, reduces the TC. Our objective in the consistent investment process is to target a high TC in order to improve the transparency of the portfolio construction process. With this in mind, we construct our FMPs and the target portfolio to reflect these active share constraints.

We constructed a second set of momentum and quality FMPs where the active share in each sector is restricted to 5 % of the portfolio size. Note that the quality of the value signal is quite poor in most of the sectors and so we persist with the same value FMP. We generated a new target portfolio and implied alpha by combining these modified momentum and quality FMPs and the original value FMP. We ran a third backtest called *BaseCFMP* (CFMP = Constrained FMP), where the active share constraints are used in the FMP generation (phase 1) of the consistent investment process rather than the final strategy. The results are summarized in Table 11.7. Note that IR has further increased to 1.01. Next we ran a backtest called *BaseCTP* (CTP = Constrained Target Portfolio), where the active share constraints are used in the target portfolio construction (phase 2) of the consistent investment process. The results are summarized in Table 11.8. This further increases the IR to 1.10. Note that the only difference between Base + CAS, BaseCFMP, and BaseCTP is that the active share constraints are imposed at different stages of the consistent investment process. Each approach incrementally improved the IR of the portfolio.

We then ran three more backtests *BaseCTP + CAS*, where the active share constraints are imposed both in the target and final portfolios; *Base CFMPTP* (CFMPTP = Constrained FMP and Target Portfolio), where the active share constraints are imposed in the momentum and quality FMPs and the final portfolio; and finally *Base CFMPTP + CAS*, where the active share constraints are imposed in all the stages of the consistent investment process, i.e., in the generation of the momentum and quality FMPs, the target portfolio, as well as the final portfolio. The results are summarized in Table 11.9. We note that adding the active share constraints to the final portfolio ensures that the final portfolio also satisfies these constraints. Moreover, the addition of these constraints also improves the IR for the BaseCFTP and BaseCFMPTP backtests.

Table 11.8 BaseCFMP backtest summary

	BaseCTP
Annualized active return	3.29 %
Avg annualized active risk	2.99 %
Avg turnover	14.89 %
Information ratio	1.10

Table 11.9 BaseCFTP + CAS, Base CFMPTP, and BaseCFMPTP + CAS summaries

	BaseCTP + CAS	BaseCFMPTP	BaseCFMPTP + CAS
Annualized active return	3.29 %	3.37 %	3.32 %
Avg annualized active risk	2.84 %	3.08 %	2.94 %
Avg turnover	14.89 %	14.89 %	14.89 %
Information ratio	1.15	1.09	1.13

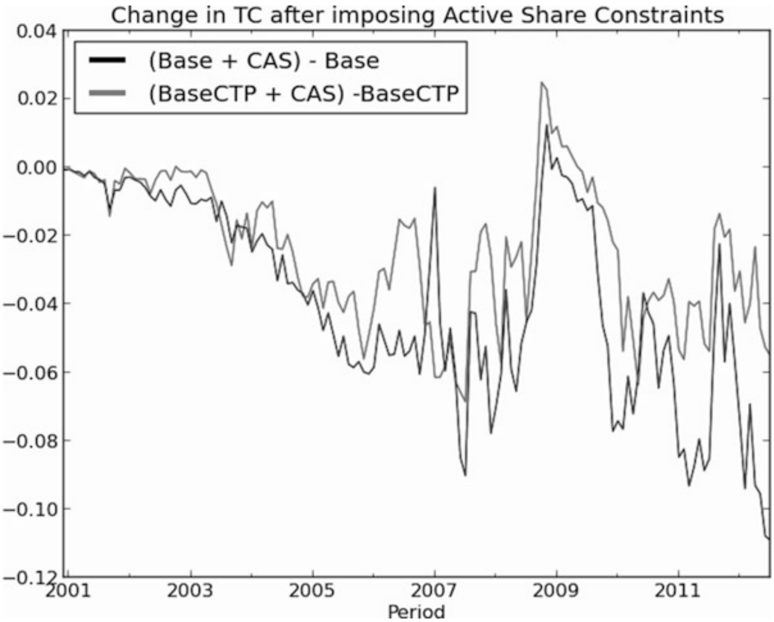


Fig. 11.2 Change in the TCs after the addition of the Active Share constraints

We conclude this section with Fig. 11.2 that compares the differences in the TCs between the (BaseCTP + CAS) and BaseCTP backtests with the corresponding differences between the (Base + CAS) and Base backtests. Note that there is less deterioration in the TC when one adds the active share constraints in the final portfolio problem in the BaseCTP backtest. This highlights our comment that it is useful to consider some of the final portfolio constraints in the construction of the implied alpha signal. Ensuring that the target portfolio also satisfies these constraints helps limit the deterioration of the TC in the realistic portfolio construction problem.

11.5 Conclusions

We presented a use case for the consistent investment approach presented in Stubbs (2013). The process breaks the original portfolio construction into three portfolio optimization problems that

1. Transform each alpha signal in the alpha factor into a factor mimicking portfolio (FMP) that represents the signal. Some of the signal specific constraints in the final portfolio problem can be added to the optimization problem that generates the FMP.
2. Combine the factor mimicking portfolios into a target portfolio by solving an MVO optimization problem. Some of the constraints in the final portfolio problem can be added to the optimization problem that generates the target portfolio.
3. Solve the actual portfolio construction problem with two changes: The first is a new alpha signal that is constructed as the implied alpha of the target portfolio and the use of a custom risk model. The second is that this portfolio problem only has the relevant implementation constraints. The final stage of the consistent process generates a portfolio that satisfies the relevant implementation constraints, and is also close to the target portfolio that is constructed in Step 2.

References

Asness, C. (1997). The interaction between value and momentum strategies. *Financial Analysts Journal, 53*(2), 29.

Black, F., & Litterman, R. (1990). *Asset allocation: combining investor views with market equilibrium*. Technical report, Goldman Sachs Fixed Income Research, 1990.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal, 48*, 28–43.

Clarke, R., de Silva, H., & Thorley, S. (2006). The fundamental law of active portfolio management. *Journal of Investment Management, 4*(3), 54–72.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465.

Gerard, X., Guido, R., & Wesselius, P. (2012). *Integrated alpha modelling*. Technical report, Advanced Risk & Portfolio Management paper, January 2012.

Grinold, R. C., & Kahn, R. N. (2000). *Active portfolio management* (2nd ed.). New York, NY: McGraw-Hill.

Jensen, M. J. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance, 23*(2), 389–416.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Markowitz, H. (1991). *Portfolio selection: efficient diversification of investments* (2nd ed.). New York: Wiley.

Qian, E., Sorenson, E. H., & Hua, R. (2007). Information horizon, portfolio turnover, and optimal alpha models. *The Journal of Portfolio Management, 34*(1), 27–40.

Sefton, J., Bulsing, M., & Scowcroft, A. (2004). *Understanding forecasting: A unified framework for combining both analysts and strategy forecasts*. Technical report, UBS, 2004.

Sivaramakrishnan, K., & Stubbs, R. A. (2013). Improving the investment process with a custom risk model: A case study with the GLER model. *The Journal of Investing, 22*(4), 129–147.
Stubbs, R. A. (2013). *Consistent portfolio management: Alpha construction*. Technical report, Axioma Inc. Research Report #044, January 2013.
Treynor, J. L. (1966). How to rate management investment funds. *Harvard Business Review, 43*(1), 63–75.
Treynor, J. L., & Black, F. (1973). How to use security analysis to improve portfolio selection. *The Journal of Business, 46*(1), 66–86.

A Technical Appendix

Proposition 1. *The FMP (11.5) with only the industry neutrality constraints carries out the equal and market cap weighted industry purifications when $W = I$ and $W = M^{-1}$ where M is the diagonal matrix whose entries are the asset market-caps, respectively.*

Proof. Consider the FMP problem

$$\begin{aligned} \min_h \ & \tfrac{1}{2} h^T W h \\ \text{s.t.} \quad & X_I^T h = 0, \\ & \alpha^T h = 1 \end{aligned} \tag{11.13}$$

that is neutral to the industry factors and where α is our alpha signal. In this case, the optimal portfolio h^* has the following expression

$$h^* = \theta W^{-1/2} \left(I - W^{-1/2} X_I \left(X_I^T W^{-1} X_I \right)^{-1} X_I^T W^{-1/2} \right) W^{-1/2} \alpha \tag{11.14}$$

where

$$\theta = \alpha^T W^{-1/2} \left(I - W^{-1/2} X_I \left(X_I^T W^{-1} X_I \right)^{-1} X_I^T W^{-1/2} \right) W^{-1/2} \alpha \tag{11.15}$$

is a positive constant. The equal and the market-cap weighted industry purifications update the α as

$$\bar{\alpha} = (I - X_I (X_I^T W^{-1} X_I)^{-1} X_I^T W^{-1}) \alpha, \tag{11.16}$$

where $W = I$ and $W = M^{-1}$, respectively. Consider the unconstrained MVO problem

$$\min_h \ \bar{\alpha}^T h - \tfrac{1}{2\theta} h^T W h \tag{11.17}$$

with the industry purified alpha, where θ is given by (11.15). The solution to this problem is given by

$$\begin{aligned} h^{MVO} &= \theta W^{-1}\bar{\alpha} \\ &= \theta W^{-1}\left(I - X_I\left(X_I^T W^{-1} X_I\right)^{-1} X_I^T W^{-1}\right)\alpha \\ &= \theta W^{-1/2}\left(I - W^{-1/2} X_I\left(X_I^T W^{-1} X_I\right)^{-1} X_I^T W^{-1/2}\right) W^{-1/2}\alpha. \end{aligned} \tag{11.18}$$

Note that h^* in (11.14) is identical to the h^{MVO} in (11.18). In other words, the optimization problem (11.13) is implicitly neutralizing the alpha signal over the industries using the weighted projection in (11.16).

Chapter 12
Empirical Analysis of Market Connectedness as a Risk Factor for Explaining Expected Stock Returns

Shijie Deng, Min Sim, and Xiaoming Huo

12.1 Introduction

Analyzing financial asset returns by identifying market-wide risk drivers and common firm-level characteristics that contribute to the explanation of expected asset returns has evolved into one major research field in the development of the modern asset pricing theory. The Capital Asset Pricing Model (CAPM) developed by Treynor (1962, 1961, Market value, time, and risk, "unpublished"),[1] Sharpe (1964), Lintner (1965a,b), and Mossin (1966) initiated this strand of research, which is referred to as the single-factor model. The single-factor model identifies a single index, or a market portfolio, as the sole driver of the return of financial assets and decomposes individual asset return risk into systematic and idiosyncratic components.

Empirical studies based on the single-factor model report mixed findings in validating CAPM as a positive economic model. Early studies such as Black et al. (1972) and Fama and MacBeth (1973) find evidence supporting a linear relationship between the average asset returns. A quantity measuring how asset returns covary with the return of market portfolio, termed market beta, is found when the data period is long. However, subsequent studies such as Fama and French (1992) and Davis (1994) provide only weak evidence in supporting CAPM. Roll (1977) points out that the CAPM cannot be empirically tested conclusively because of the difficulty in measuring the risk-return characteristics of the market portfolio.

[1]See French (2003) for details of these references.

S. Deng (✉) • M. Sim • X. Huo
Steward School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: deng@gatech.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_12

Roll and Ross (1994), Kandel and Stambaugh (1995) demonstrate that even very small deviations from the efficient portfolio can yield the linear relationship between risk and expected returns insignificant.

The inconclusive empirical testing results of the single-factor model combined with evidence on that firm-level fundamental variables such as earnings-to-price (E/P) ratio (Basu (1977)) predict higher asset returns than those predicted by market beta prompt that the market beta alone may not be able to explain the cross-sectional variation in the asset returns. This leads to a growing research literature on extending the single-factor CAPM to a multi-factor model by using firm-level fundamental variables such as E/P ratio (Basu (1977)) and market value to book value ratio (e.g., Rosenberg et al. (1985), De Bondt and Thaler (1987)), market-level variables such as price momentum (Jegadeesh and Titman (1993)), and macroeconomic variables such as trading liquidity (Paster and Stambaugh (2003)) to explain the expected asset returns.

Even with the extended multi-factor models such as Rosenberg (1974), Roll and Ross (1980), and Fama and French (1993), empirical studies with equity market data still do not generate clear-cut positive results. Researchers are constantly searching for alternative risk factors that may have stronger explanatory variables for the cross-sectional asset returns. Paster and Stambaugh (2003) is one such example in which the authors show that individual stock's sensitivity to market level liquidity innovation can be a significant driver for asset return variations. In a similar vein, Sim et al. (2014) show that individual stock's sensitivity to the overall market connectedness forms a promising risk factor that helps to explain the expected stock returns.

In this article, we introduce a quantitative measure for the connectedness of financial markets as proposed in Sim et al. (2014). We demonstrate via empirical tests using a two-factor model that the market connectedness measure holds explanatory power the expected stock returns. The remainder of this article is organized as follows. Section 12.2 presents the classical approaches for empirically testing CAPM and its multi-factor extensions. The description of the alternative measure for market connectedness and its construction are given in Sect. 12.3. Empirical tests on whether the market connectedness corresponds to a new source of systematic risk driving the stock returns are performed in Sect. 12.4. Finally, we present results and conclude in Sect. 12.5.

12.2 CAPM and the Multi-Factor Asset Pricing Model

Let R_s, R_M, R_f denote the returns of an asset s, the market portfolio M, and the risk-free asset, respectively. CAPM specifies a linear relationship between the return of any individual financial asset and that of a market portfolio. Namely,

$$E[R_s] = R_f + \beta_{s,M}(E[R_M] - R_f) \tag{12.1}$$

where $E[\cdot]$ denotes expected value, and $\beta_{s,M} \equiv \frac{Cov(R_s,R_M)}{\sigma_M^2} = \rho_{s,M}\frac{\sigma_s}{\sigma_M}$ is the market beta of asset s measuring the systematic risk exposure of the excess return of s to the risk of the market portfolio M. In the definition of $\beta_{s,M}$, σ_R and $\rho_{s,M}$ denote, respectively, the volatility of R and the correlation coefficient between the returns R_s and R_M.

12.2.1 Empirical Testing of CAPM

Equation (12.1) is often referred to as the Security Market Line (SML) and it leads to the usual form of the testing hypothesis of the empirical investigation of CAPM.

As CAPM is a single-period *ex ante* model and asset returns are not known in *ex ante*, researchers use *ex post* returns to test CAPM instead. Specifically, market beta of asset s is estimated through the following equation using historical data:

$$r_{s,t} - r_{f,t} = a_s + \beta_{s,M}(r_{M,t} - r_{f,t}) + \epsilon_{s,t}, \tag{12.2}$$

where in each period a_s is a constant return earned by asset s, $r_{s,t}$ is the return of asset s at time t, $r_{f,t}$ is the risk-free rate at time t, and $\epsilon_{s,t}$ is the noise in the realized return of s. The estimated market beta $\beta_{s,M}$ is used as explanatory variable to test the cross-sectional equation (12.3).

$$r_{s,t} = \alpha_0 + \alpha_1\beta_{s,M} + \eta_{s,t}, \tag{12.3}$$

where α_0 is the expected return of a risk-free asset (or, a zero-beta portfolio), α_1 is the expected excess return of the market portfolio (or, the market risk premium), and $\eta_{s,t}$ is the noise term. If the cross-sectional test yields a statistically significant value of α_1, then the validity of CAPM is supported.

While initial empirical research such as Black et al. (1972), Fama and MacBeth (1973) found supporting evidence of high beta assets tend to generate high level of returns that is consistent with the linear relationship in (12.2), later research working with a larger amount of historical data (e.g., Fama and French (1992), Davis (1994)) found that the empirical evidence is rather weak. Further evidence on the market portfolio falling short in fully explaining asset returns such as Basu (1977), Banz (1981), Rosenberg et al. (1985), Jegadeesh and Titman (1993) sparks a vast body of research on extending the CAPM model, in the spirit of the arbitrage-pricing model of Ross (1976), to a multi-factor model as proposed by Fama and French (1993, 1996), Carhart (1997), Frankel and Lee (1998), Paster and Stambaugh (2003), among others.

12.2.2 Multi-Factor Asset Return Model

In a general multi-factor asset return model, the excess return of asset s (namely, the amount in excess to the risk-free return rate r_f), denoted by $\bar{r}_s$, is attributed to its exposure to a set of N_c non-diversified systematic risk factors. Specifically, we have

$$\bar{r}_s = \alpha + \sum_{c=1}^{N_c} \beta_{s,c}\bar{f}_c + \epsilon_s, \tag{12.4}$$

where $\bar{f}_c$ is the excess return of the c^{th} systematic risk factor, ϵ_s is the asset-specific residual after removing the impact of all factors from the excess return of asset s. It represents the diversifiable risk that is specific to asset s, and $\beta_{s,c}$ measures the exposure of the excess return of asset s to the systematic risk-factor c and is termed *factor beta*.

Various types of observable variables have been proposed as alternative systematic risk factors. These include firm-level variables (e.g., earnings-to-price, book-to-market, market capitalization level), market-level variables such as price momentum, and macro-economics level variables such as liquidity. There are also pure statistical factors obtained through analyzing the covariance matrix of asset returns directly (see Connor and Korajczyk (2010) for an extensive review of the risk factor models).

12.3 Market-Connectedness and Systematic Risk in Asset Returns

As the scope of financial markets has been expanded tremendously over the recent decades through introductions of vast amounts of stocks and diverse derivatives products, a question arises as to whether this yields a more expanded investment opportunity set for investors in general.

Anecdotal evidences indicate that the levels of interactions within and among financial markets have increased significantly over the last decade. Thus the expanding landscape of financial markets may not result in a much expanded investment opportunity set. In fact, markets with highly correlated traded assets, even with the total market capitalization being large, do not necessarily provide diverse investment opportunities to market participants. Market participants must comprehend the inter-related structures in markets in order to truly assess the investment opportunity set so that they can practice portfolio diversification and risk managements effectively.

12.3.1 Alternative Measures for Financial Market Connectedness

The research strand on the study of market connectedness[2] has been growing and focusing on quantifying the level of association in financial assets in order to assess overall market structures from the perspective of a graph or a network. For example, Billio et al. (2012) and Diebold and Yilmaz (2011) construct their connectedness measures in financial institutions to measure the level of systemic risk during the global recession period in 2007–2008 and provide empirical evidence that their measures are related to the cycle of economy.

In analyzing the connectedness of financial markets and its impact on the investment opportunity set, Sim et al. (2014) propose a market connectedness measure, termed modularity, to quantify the level of connectedness of financial markets. They take a different approach to quantify the market connectedness through analyzing *the clustering tendency* in stock markets.

According to recent studies, the cluster property, where entities with similar characteristics tend to form a subgroup or a cell, is one of the most evident and important structural properties in financial markets. Materassi and Innocenti (2009) provide empirical evidence that the major stocks in the US can be drawn in tree structure, a special case of cluster structure, where branches of the tree connect the highly correlated stocks together. Pojarliev and Levich (2010) classify foreign exchange investing funds into two groups and proposed a few crowdedness measures for co-movement tendency of market participants. Chandrasekaran et al. (2012) conduct an empirical analysis on the US stock market through a hidden Gaussian graphical model, which shows that the clustering tendency across the universe of stocks is observable even after eliminating a few common drivers of the stock market.

Following the clustering studies in Materassi and Innocenti (2009) and Chandrasekaran et al. (2012), Sim et al. (2014) classify the correlation elements in stock markets into two groups: one group containing stocks that tend to be highly correlated with each other and the other group containing stocks that fluctuate along with the flucturation of the cycle of economy. Connectedness measures are then constructed by measuring the relative difference between the respective average correlations of the two groups. The relationship between the proposed connectedness measure and the movements of individual stocks are further explored.

[2]The development of this line of study is largely grounded in the development of graph theory or network theory that are actively studied in the various disciplines such as combinatorics, computer science, physics, and (bio)-statistics.

12.3.2 Market Connectedness Measure: Modularity

This section offers a bottom-to-top approach for constructing the modularity measure using Pearson's pairwise correlation. Let $C(i,j)$ denote Pearson's pairwise correlation between two stock returns. Namely,

$$C(i,j) \equiv \rho_{i,j} = \frac{Cov(r_i, r_j)}{std.dev.(r_i)std.dev.(r_j)} \tag{12.5}$$

where r_i and r_j are the returns of stock i and j, respectively. Using pairwise correlations as a building block, the *connectedness between two groups of stocks* is defined as follows:

$$C(A,B) := \text{Mean}\left(\{C(i,j) | \forall i \in A, \forall j \in B, i \neq j\}\right) \tag{12.6}$$

where A and B are two groups of stocks, and $Mean(\cdot)$ calculates the mean of elements in a set. Note that the groups A and B are allowed to have overlaps (or, even be identical to each other). The condition $i \neq j$ excludes trivial self-correlations for overlapping stocks.

Let V denote the universe of stocks considered for investment in the market. A *partition* P of V is defined as $P = \{V_1, V_2, \ldots, V_k\}$, where $V = \bigcup_{c=1}^{k} V_c$, $V_i \cap V_j = \emptyset$, and V_i denotes the i^{th} sub-group of stocks. Each sub-group is termed a *cell*. *Clustering* or *cluster analysis* on correlation matrix is a task of finding the best partition P for V such that the pairwise correlation of returns between stocks within each cell are generally higher than the return correlations of stocks that belong to different cells.

The connectedness of stock returns in universe V is defined with respect to a given partition P. Specifically, the *inner-sector connectedness (INSC)* is the average of all pairwise correlations within the cells in the partition P,

$$\text{INSC}(P) := \text{Mean}\left(\bigcup_{c=1}^{k} \{C(i,j) | (i,j) \in (V_c, V_c), i \neq j\}\right). \tag{12.7}$$

Similarly, the *inter-sector connectedness (ITSC)* is defined as the average of all correlations across the cells in the partition P. Namely,

$$\text{ITSC}(P) := \text{Mean}\left(\bigcup_{c_1=1}^{k-1} \bigcup_{c_2=c_1+1}^{k} \{C(i,j) | (i,j) \in (V_{c_1}, V_{c_2})\}\right). \tag{12.8}$$

If the partition P exhibits a prominent cluster structure in the asset return correlation matrix, then INSC(P) is expected to be much higher than ITSC(P), meaning that the returns of assets in each cell of P are much more dependent on each other in their own cell than they are with the returns of stocks in other cells. Therefore, a high

INSC(P) in conjunction with a low ITSC(P) implies that the partition P represents a very prominent clustering of asset returns in universe V while a low INSC(P) combined with a high ITSC(P) implies the contrary.

The modularity of the asset returns in universe V with respect to the partition P is defined as the difference between INSC(P) and ITSC(P) with an intuitive meaning of capturing the significance of a clustering structure represented by partition P. Namely,

$$\mathrm{MOD}(P) := \mathrm{INSC}(P) - \mathrm{ITSC}(P). \tag{12.9}$$

Clearly, identifying a clustering structure among the myriad of financial assets traded in the markets, if such a structure exists, is a first important step towards the understanding of connectedness of various financial markets. Sim et al. (2014) adopt the *Modulated Modularity Clustering (MMC)* method proposed by Stone and Ayroles (2009) for detecting the clustering structure in the financial security returns.

12.4 Modularity Index as a Systematic Risk Factor: Empirical Analysis

In this section, we take 60 major stocks from the top of the Fortune 500 U.S. firms, ranked by their operating revenue in 2011, as the universe of investment opportunities and report the clustering structure obtained by applying the MMC method. Pairwise return correlations are calculated based on daily return data from the The Center for Research in Security Prices (CRSP) Database provided by Wharton Research Data Services (WRDS).

Using the identified clusters as a partition, the market connectedness measures and the modularity index are computed. We demonstrate that one may construct portfolios using US equities based on their sorted beta with respect to the modularity index and generate excess returns, which is greater than what is predicted by a CAPM model.

12.4.1 Clusters of Asset Returns over a Long Period

The pairwise return correlations of the 60 stocks (full-list given in Table 12.1) are computed using the daily close price in CRSP from 1/1/2002 to 12/31/2011 to form a sample correlation matrix. Twelve clusters (or, cells) are obtained after applying the MMC algorithm to this sample correlation matrix. The identified clusters are given in Fig. 12.1. Note that the partitioning clusters identified by the MMC method

Table 12.1 List of 60 stocks chosen based on 2011 Fortune 500 ranking

Cell	Ticker	Company name	Revenues ($ millions) (in 2011)	SIC code	SIC description
1	CVS	CVS Caremark	107,750	5912	Retail-Drug Stores and Proprietary Stores
	WAG	Walgreen	72,184	5912	Retail-Drug Stores and Proprietary Stores
	SYY	Sysco	39,324	5141	Wholesale-Groceries, General Line (merchandise)
	ESRX	Express Scripts Holding	46, 128	8093	Services-Specialty Outpatient Facilities, NEC
	KFT	Kraft Foods	54,365	2052	Cookies & Crackers
	WLP	WellPoint	60,711	6324	Hospital & Medical Service Plans
2	KO	Coca-Cola	46,542	2086	Bottled & Canned Soft Drinks & Carbonated Waters
	PEP	PepsiCo	66,504	2086	Bottled & Canned Soft Drinks & Carbonated Waters
	PG	Procter & Gamble	82,559	2841	Soap and Other Detergents
	PFE	Pfizer	67,932	2834	Pharmaceutical Preparations
	JNJ	Johnson & Johnson	65,030	2834	Pharmaceutical Preparations
	MRK	Merck	48,047	2834	Pharmaceutical Preparations
3	GE	General Electric	147,616	3511	Turbines and Turbine Generator Sets
	DOW	Dow Chemical	59,985	2821	Plastic Materials, Synth Resins & Nonvulcan Elastomers
	JCI	Johnson Controls	40,833	2531	Public Bldg & Related Furniture
	MET	MetLife	70,641	6311	Life Insurance
	PRU	Prudential Financial	49,045	6311	Life Insurance
4	XOM	Exxon Mobil	452,926	2911	Petroleum Refining
	COP	ConocoPhillips	237,272	2911	Petroleum Refining
	CVX	Chevron	245,621	2911	Petroleum Refining
	MRO	Marathon Petroleum	73,645	2911	Petroleum Refining
	CAT	Caterpillar	60,138	3531	Construction Machinery & Equip

Cell	Ticker	Company name	Revenues ($ millions) (in 2011)	SIC code	SIC description
5	MSFT	Microsoft	69,943	7370	Services-Computer Programming, Data Processing, Etc.
	DELL	Dell	62,071	3570	Computer & office Equipment
	IBM	International Business Machines	106, 916	3571	Electronic Computers
	AAPL	Apple	108,249	3571	Electronic Computers
	UTX	United Technologies	58, 190	3724	Aircraft Engines & Engine Parts
	HPQ	Hewlett-Packard	127,245	3571	Electronic Computers
	INTC	Intel	53,999	3679	Electronic Components, NEC
	CSCO	Cisco Systems	43,218	3674	Semiconductors & Related Devices
	INTL	INTL FCStone	75,498	6211	Security Brokers, Dealers & Flotation Companies
	AMZN	Amazon.com	48,077	7370	Services-Computer Programming, Data Processing, Etc.
6	CAH	Cardinal Health	102,644	5122	Wholesale-Drugs, Proprietaries & Druggists' Sundries
	MCK	McKesson	112,084	5122	Wholesale-Drugs, Proprietaries & Druggists' Sundries
	ABC	AmerisourceBergen	80,218	5122	Wholesale-Drugs, Proprietaries & Druggists' Sundries
	UNH	UnitedHealth Group	101,862	6324	Hospital & Medical Service Plans
7	WFC	Wells Fargo	87,597	6021	National Commercial Banks
	JPM	J.P. Morgan Chase & Co.	110,838	6712	Bank Holding Companies
	BAC	Bank of America Corp.	115,074	6021	National Commercial Banks
	AIG	American International Group	71,730	6331	Fire, Marine & Casualty Insurance
	C	Citigroup	102,939	6021	National Commercial Banks
8	SUN	Sunoco	45,765	2911	Petroleum Refining
	VLO	Valero Energy	125,095	2911	Petroleum Refining
	EPD	Enterprise Products Partners	44,313	4922	Natural Gas Transmission

(continued)

Table 12.1 (continued)

9	TGT	Target	69,865	5331	Retail-Department Stores
	WMT	Wal-Mart Stores	446,950	5331	Retail-Department Stores
	LOW	Lowe's	50,208	5211	Retail-Lumber & Other Building Materials Dealers
	HD	Home Depot	70,395	5211	Retail-Lumber & Other Building Materials Dealers
	BBY	Best Buy	50,272	5731	Retail-Radio, TV & Consumer Electronics Stores
	COST	Costco Wholesale	88,915	5330	Food & Drug Retailers
	UPS	United Parcel Service	53,105	4215	Courier Services, Except By Air industry
10	ADM	Archer Daniels Midland	80,676	2075	Soybean Oil Mills
	BA	Boeing	68,735	3721	Aircraft
	LMT	Lockheed Martin	46,692	3764	Space Propulsion Units and Parts
11	F	Ford Motor	136,264	3711	Motor Vehicles & Passenger Car Bodies
	DIS	Walt Disney	40,893	7996	Services-Amusement Parks
	VZ	Verizon Communications	110,875	4813	Telephone Communications (No Radiotelephone)
	T	AT&T	126,723	4813	Telephone Communications (No Radiotelephone)
12	KR	Kroger	90,374	5411	Retail-Grocery Stores
	SWY	Safeway	43,630	5411	Retail-Grocery Stores

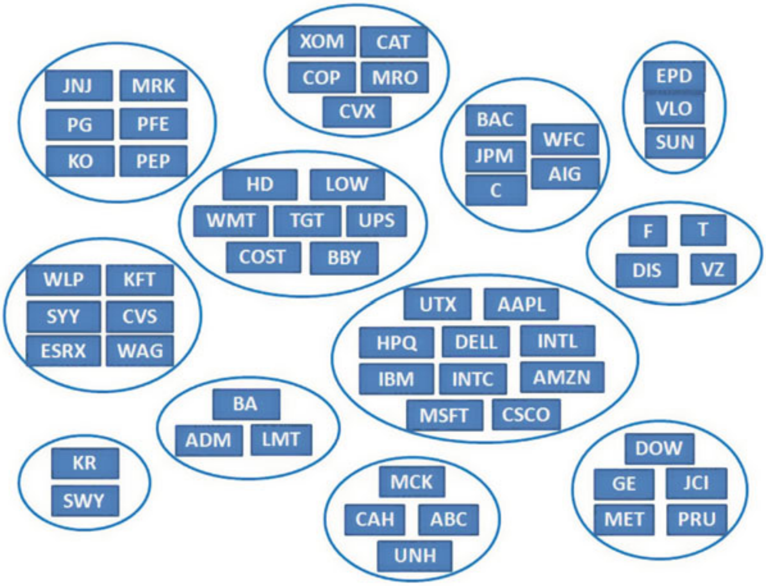


Fig. 12.1 Cluster analysis with 10 years returns of 60 major stocks by MMC algorithm

do not exactly match with those categorized by the Standard Industry Classification (SIC) codes (see Table 12.1). Based on this cluster partition, a modularity index, MOD, is constructed from Eqs. (12.7), (12.8) and (12.9).

12.4.2 Modularity: A Systematic Risk Factor

To empirically test whether the modularity index MOD is a valid systematic risk factor, we conjecture that the decile portfolio sorted by the individual stock's sensitivity to MOD would show significant differences in return. We follow a similar regression procedure as that used by Fama and MacBeth (1973) to estimate the beta of each stock with respect to MOD_t over time based on a two-factor model (namely, setting $N_c = 2$ in model (12.4)). The detailed steps of the procedure are described in Sim et al. (2014). After getting the beta-to-MOD of each stock, we construct the decile portfolios based on sorted values of beta-to-MOD, where the top decile portfolio consists of stocks with the least beta-to-MOD values.

Using data from the period of January 1992 to December 2011, Table 12.2 shows that beta-to-MOD sorted decile portfolio creates systematic return difference that are not explained by CAPM. The first row presents the annualized return and standard deviation of each decile portfolio. The last column corresponds to the difference of 1-portfolio and 10-portfolio that is equivalent to the net zero investment portfolio where investors buy the first decile and sell the last decile in

Table 12.2 Returns and alphas of the decile portfolios (January 1992–December 2011)

	1	2	3	4	5	6	7	8	9	10	'1–10'
Return (p.a.)	13.99	10.23	10.29	11.41	10.06	8.21	8.48	11.11	7.72	6.4	7.59
s.d. (p.a.)	20.94	17.1	14.29	14.25	14.45	14.78	14.99	15.67	18.5	20.9	14.8
CAPM alpha	4.25	1.43	2.51	3.54	2	0.25	0.35	2.6	−1.69	−3.66	7.92
(t-statistics)	(1.66)	(0.79)	(1.58)	(2.39)	(1.46)	(0.15)	(0.23)	(1.84)	(−0.94)	(−1.66)	(2.38)

Fig. 12.2 Cumulative return on enhancement scenarios with MOD_P

Table 12.3 Performance of enhanced market index portfolio (January 1992–December 2011)

	MKT	MKT+10 % MOD_s	MKT+20 % MOD_s
Return (p.a.)	8.94 %	9.69 %	10.45 %
Std (p.a.)	15.46 %	15.44 %	15.57 %
SR (monthly)	0.109	0.123	0.136

the same amount. The second row present the level of alphas (and the t-statistics for the null hypothesis of alphas being zero) when fitting the excess returns of the decile portfolios to CAPM. Clearly, the return of the 1–10 portfolio reported in Table 12.2 cannot be fully attributed to that of the market portfolio in CAPM.

Indeed, the 1–10 portfolio based on the MOD factor enlarges the investment opportunity set for investors. Figure 12.2 illustrates that, if one adds different weights, such as 10 % and 20 %, of the 1–10 portfolio to the proxy market portfolio, then the resulting overall portfolio outperforms the proxy market portfolio. The corresponding Sharpe ratios are higher than that of the proxy market portfolio (see the last row of Table 12.3).

12.5 Conclusion

In this article, a quantitative measure for quantifying the connectedness of financial markets is briefly introduced. Through empirical tests, we demonstrate that this alternative measure for market connectedness, termed modularity in Sim et al. (2014), can act as a new risk factor for explaining the expected stock returns under the multi-factor asset pricing model framework. Using the U.S. equity market data from 1992 to 2011, decile portfolio analysis based on the beta-to-modularity indeed generates significant excess returns that cannot be explained by CAPM. Empirical tests also reveal that the properly constructed decile portfolios based on asset return sensitivity to market connectedness enhances investment performance by enlarging the existing investment opportunity set.

Acknowledgements We thank Dr. John Guerard for constructive comments on this work. Any remaining errors are our own. The third author is supported by (while serving at) the National Science Foundation. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Banz, W. R. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3–18.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. *Journal of Finance, 12*(3), 129–156.

Billio, M., Getmansky, M., Lo, W. A., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics, 104*(3), 535–559.

Black, F., Jensen, M., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. Jensen (Ed.), *Studies in the theory of capital markets* (pp. 79–121). New York: Praeger.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82.

Chandrasekaran, V., Parrilo, P. A., & Willsky, A. S. (2012). Latent variable graphical model selection via convex optimization. *Annals of Statistics, 40*(4), 1935–1967.

Connor, G., & Korajczyk, R. A. (2010). Factor models in portfolio and asset pricing theory. In: *Handbook of portfolio construction* (pp. 401–418). New York: Springer.

Davis, J. (1994). The cross-section of realised stock returns: the pre-compustat evidence. *Journal of Finance, 49*, 1579–1593.

De Bondt, W., & Thaler, R. (1987). Further evidence of investor overreaction and stock market seasonality. *Journal of Finance, 42*, 557–581.

Diebold, F., & Yilmaz, K. (2011). *On the network topology of variance decompositions: Measuring the connectedness of financial firms*. Working Paper.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.

Fama, E., & MacBeth, J. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy, 81*(3), 607–636.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*(1), 55–84.

Frankel, R., & Lee, M. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics, 25*(3), 283–319.

French, C. W. (2003). The treynor capital asset pricing model. *Journal of Investment Management, 1*(2), 60–72.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*(1), 65–91.

Kandel, S., & Stambaugh, R. (1995). Portfolio inefficiency and the cross-section of expected returns. *Journal of Finance, 50*, 157–184.

Lintner, J. (1965a). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*(1), 13–37.

Lintner, J. (1965b). Securities prices, risk, and maximal gains from diversification. *Journal of Finance, 20*(4), 587–615.

Materassi, D., & Innocenti, G. (2009). Unveiling the connectivity structure of financial networks via high-frequency analysis. *Physica A: Statistical Mechanics and Its Applications, 388*(18), 3866–3878.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768–783.

Paster, L., & Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*(3), 642–685.

Pojarliev, M., & Levich, R. (2010). *Detecting crowded trades in currency funds*. Working Paper.

Roll, R. (1977). A critique of the asset pricing theory's tests; part i: on past and potential testability of the theory. *Journal of Financial Economics, 4*, 129–176.

Roll, R., & Ross, S. A. (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance, 35*(5), 1073–1103.

Roll, R., & Ross, S. A. (1994). On the cross-sectional relation between expected returns and betas. *Journal of Finance, 49*, 101–121.

Rosenberg, B. (1974). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis, 9*, 263–274.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management, 11*(Spring), 9–17.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341–360.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425–442.

Sim, M., Deng, S., & Huo, X. (2014). *Market connectedness and expected stock returns*. Working paper, Georgia Institute of Technology.

Stone, E., & Ayroles, J. (2009). Modulated modularity clustering as an exploratory tool for functional genomic inference. *PLoS Genet, 5*, e1000479.

Treynor, J. L. (1962). Toward a theory of market value of risky assets. Unpublished manuscript. In R. Korajczyk (Ed.), A later version was published in 1999 as a chapter in *Asset pricing and portfolio performance* (pp. 15–22). London: Risk Books.

Chapter 13
The Behaviour of Sentiment-Induced Share Returns: Measurement When Fundamentals Are Observable

Richard A. Brealey, Ian A. Cooper, and Evi Kaplanis

JEL classification: G11, G12, G14

In this article we investigate the relationships between investor sentiment and deviations of share prices from fundamental values. To do this we use a sample of shares for which a large part of the fundamental value is observable: upstream oil stocks. We measure their fundamental values using oil and gas prices and the forward oil price contango.

We focus on upstream oil stocks because there is a direct relationship between the present value of these stocks and the oil price. In a world where output prices minus extraction costs obey the Hotelling Principle, the value of natural resource companies depends only on current prices less extraction costs. Miller and Upton (1985a, 1985b) test this proposition and find that it provides a good explanation of the variation in value of a sample of oil producers. Hence a large part of the fundamental value of upstream oil stocks is observable. We make use of this present value condition to split the return on our sample into the part that represents fundamentals and the part that is a deviation from fundamentals. The attraction of the Hotelling Principle is that it avoids the need to forecast future cash flows and to estimate discount rates. A more general and less parsimonious model might include additional variables, such as exchange rates and proxies for the discount rate.

Following Baker–Wurgler (2006), we test the impact of sentiment using a portfolio that is long high-variance stocks and short low-variance stocks (the "Hi-Lo" portfolio). We find that two types of sentiment predict returns: retail sentiment, which predicts momentum, and the Baker–Wurgler (2006) measure of sentiment,

R.A. Brealey • I.A. Cooper (✉) • E. Kaplanis
London Business School, Sussex Place, Regent's Park, London, NW1 4SA, UK
e-mail: icooper@london.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_13

which predicts reversion to fundamentals. We find that the influence of sentiment in each case is time-varying. In particular, the ability of sentiment to predict returns appears only after 2000. Contrary to the theory that sentiment mainly affects the deviations from fundamental value of hard-to-arbitrage stocks, we find that both measures of sentiment influence prices through the fundamentals themselves rather than through deviations from fundamentals.

If the Hi and Lo portfolios had similar loadings on fundamental variables, the net Hi-Lo portfolio would be hedged against fundamental effects and we should not observe fundamentals affecting the returns on this portfolio. However, in our data the loadings on fundamental factors of the Hi and Lo portfolios are different, so the combination of a long and short position in this portfolio does not eliminate its exposure to fundamentals. Methodologically, this raises the issue that tests based on such portfolios do not avoid the need to control for fundamentals.

The remainder of the chapter is organised as follows. In Section 13.1 we give a brief review of related literature. In Sect. 13.2 we develop our tests and in Sect. 13.3 we describe our data. Section 13.4 presents our main tests of the influence of sentiment on returns with and without controls for fundamentals. Section 13.5 provides some robustness tests. Section 13.6 concludes.

13.1 Related Literature

Our work is broadly related to a number of studies that have found evidence of serial dependence in returns. Evidence of momentum over periods of six to twelve months is provided by amongst others Lehmann (1990), Jegadeesh (1990), Jegadeesh and Titman (1993, 2001), Asness, Moskowitz, and Pedersen (2013), and Moskowitz, Ooi, and Pedersen (2011). Evidence that this medium-term momentum is followed by longer-term mean reversion comes from variance-ratio tests (Poterba & Summers, 1988, Lo & MacKinlay, 1988, Cutler, Poterba, & Summers, 1991) and autocorrelation tests (Fama & French, 1988). Evidence that high short-term variance is related to deviations from fundamentals comes from excess variance tests (Shiller, 1981, LeRoy & Porter, 1981).

Sentiment-based explanations of momentum, mean-reversion, and deviations from fundamentals envisage these effects as arising from behavioural biases by naïve investors combined with costs of arbitrage. For example, Daniel, Hirshleifer, and Subrahmanyam (1998) present a model in which a combination of overconfidence and biased self-attribution create both under- and over-reaction. Similarly, Barberis, Shleifer, and Vishny (1998) appeal to the behavioural biases of representativeness and conservatism to show how these can result in under- and over-reaction. In both papers asset prices can be decomposed into one part that reflects fundamentals and another consisting of deviations from fundamentals. The effect of sentiment on asset prices operates through the deviations from fundamentals.

Empirical evidence on the link between sentiment and returns requires a measure of sentiment. A number of suggestions have been proposed. Many of these reflect the view that sentiment changes are driven by retail investors. Possible proxies include flows into mutual funds (Brown, Goetzmann, Hiraki, Shirishi, & Watanabe, 2003), buy–sell imbalances by retail investors (Kumar & Lee, 2006), IPO volume and initial returns (Baker & Wurgler, 2006), market turnover (Baker & Wurgler, 2006), closed-end fund discounts, (Lee, Shleifer, & Thaler, 1991, Chen, Kan, & Miller, 1993, Swaminathan, 1996, and Neal & Wheatley, 1998), the growth stock premium (Baker & Wurgler, 2006), and survey data (Qiu & Welch, 2004, Brown & Cliff, 2004, 2005). These data have been used either singly or in combination as sentiment measures to test hypotheses about the relationship between sentiment and subsequent stock returns.

Our tests of the effect of sentiment are most closely related to Baker and Wurgler (2006, 2007, 2012) and Baker, Wurgler, and Yuan (2012). Baker and Wurgler divide their sample of stocks into ten portfolios on the basis of their prior volatility, which serves as a proxy for difficulty of arbitrage. They find that returns on the more volatile stocks are lower following a time of optimism, and that returns are higher following a time of pessimism. For the less volatile stocks that are easier to arbitrage the reverse is true. They develop a measure of investor sentiment which they find predicts returns for portfolios that are long the more volatile stocks and short less volatile stocks. This finding is consistent with a combination of behavioural biases and limits to arbitrage.

Barberis, Shleifer and Wurgler (2005) stress the importance of controlling for fundamentals when measuring the effect of sentiment on security prices. For example, Derrien and Kecskés (2009) show that the effect of sentiment on equity issuance disappears once controls for fundamentals are included. Baker and Wurgler's use of a Hi-Lo portfolio will be effective in controlling for fundamentals only if the long and short positions have equal loadings on fundamental factors. The alternative is to attempt to control directly for fundamentals. Brown and Cliff (2005) use as their dependent variable estimates of deviations from fundamental value based on the dividend discount model. They find that these deviations are positively related to a sentiment measure derived from survey data. However, the dividend discount model gives a very noisy observation of fundamental value. For our sample of stocks, we have a more direct measure of fundamental value than Brown and Cliff and so are able to perform a more powerful test of the way that sentiment is transmitted to share prices and returns.

13.2 Hypotheses and Tests

We test the implications of the hard-to-arbitrage hypothesis using a simple empirical procedure that relates sentiment measures to deviations of share prices from fundamentals and also to the fundamentals themselves. We split the log share price,

P_t, into a component reflecting fundamental value, F_t, and a separate component, NF_t, which is the deviation from fundamental value:

$$\mathrm{NF}_t = P_t - F_t \tag{13.1}$$

We assume that prices are affected by the actions of two types of traders. One type is an arbitrageur, whose behaviour is captured by the Baker–Wurgler sentiment measure S_t. The other type is a naïve trend-follower, whose behaviour is captured by a bullish retail sentiment measure, B_t. Fundamentals and non-fundamentals may respond to both sentiment variables:

$$F_t - F_{t-1} = \Delta F_t = \theta_{B,F} B_{t-1} + \theta_{S,F} S_{t-1} + \mathrm{e}_{Ft} \tag{13.2}$$

$$\mathrm{NF}_t - \mathrm{NF}_{t-1} = \Delta \mathrm{NF}_t = \theta_{B,\mathrm{NF}} B_{t-1} + \theta_{S,\mathrm{NF}} S_{t-1} + \mathrm{e}_{\mathrm{NF}t} \tag{13.3}$$

We model the response to Baker–Wurgler sentiment by assuming that arbitrageurs push prices down when this sentiment variable is high, making $\theta_{S,\mathrm{F}} < 0$ and $\theta_{S,\mathrm{NF}} < 0$. The effect of naïve trend-followers pushes prices up when Bullish sentiment is high, making $\theta_{B,F} > 0$ and $\theta_{B,\mathrm{NF}} > 0$.

The Baker–Wurgler sentiment variable is a measure of mispricing and so should rise when the deviation from fundamentals increases:

$$S_t = \theta_{S,S} S_{t-1} + \theta_{S,\mathrm{NF}} \Delta \mathrm{NF}_t + \mathrm{e}_{St} \tag{13.4}$$

where $\theta_{S,\mathrm{NF}} > 0$. We expect S_t to be highly persistent, reflecting the long cycle of swings in mispricing. We hypothesise that the bullish sentiment indicator reflects trend-following behaviour:

$$B_t = \theta_{B,B} B_{t-1} + \theta_{B,P} (P_t - P_{t-1}) + \mathrm{e}_{Bt} \tag{13.5}$$

$\theta_{B,P} > 0$. We expect B_t to be less persistent than S_t, reflecting shorter cycles in momentum sentiment.

We stack Eqs. (13.2)–(13.5) to form the system shown below, which we estimate using VAR. Our null hypothesis is that the sentiment variables affect deviations from fundamentals but not the fundamentals themselves. This would be consistent with the limits-to-arbitrage hypothesis, whereby sentiment causes deviations from fundamentals when such deviations are hard to arbitrage. The indicated signs of the key parameters under the null hypothesis are shown in Table 13.1. In particular, the standard hypothesis is that sentiment affects deviations from fundamentals, implying $\theta_{S,F} = 0$, $\theta_{\mathrm{B,F}} = 0$, $\theta_{S,\mathrm{NF}} < 0$, $\theta_{B,\mathrm{NF}} > 0$. As an alternative, we test the hypothesis that the influence of sentiment occurs via its effect on fundamentals, which implies that $\theta_{S,F} < 0$, $\theta_{B,F} > 0$, $\theta_{S,\mathrm{NF}} = 0$, $\theta_{B,\mathrm{NF}} = 0$.

Table 13.1 Structure of the VAR and hypotheses about coefficients

	Equation			
	(1)	(2)	(3)	(4)
	Dependent variable			
Independent variables	$F_t - F_{t-1}$	$NF_t - NF_{t-1}$	S_t	B_t
$F_{t-1} - F_{t-2}$	$\theta_{F,F} = 0$	$\theta_{F,NF} = 0$	$\theta_{F,S} = 0$	$\theta_{B,P} > 0$
$NF_{t-1} - NF_{t-2}$	$\theta_{NF,F} = 0$	$\theta_{NF,NF} = ?$	$\theta_S > 0$	$\theta_{B,P} > 0$
S_{t-1}	$\theta_{S,F} = 0$	$\theta_{S,NF} < 0$	$\theta_{S,S} >> 0$	$\theta_{S,B} = ?$
B_{t-1}	$\theta_{B,F} = 0$	$\theta_{B,NF} > 0$	$\theta_{B,B} = ?$	$\theta_{B,B} > 0$

13.3 Data

Our sample consists of the stocks of the 121 US oil exploration and production companies quoted on the NYSE during the period March 1983 to January 2011. We define an exploration or production company as one with a North American Industrial Classification (NAIC) of 211111 or a Standard Industrial Classification (SIC) of 1311. By limiting our sample in this way, we exclude refining companies that are likely to have very different loadings on the oil and gas factors.

We assume that fundamental value is a function of the month-end spot price of West Texas Intermediate oil and the spot wellhead price of West Texas natural gas. We also use the change in the contango in oil prices, where contango is measured as the log of the price of the sixth most distant futures contract less that of the price of the closest futures contract[1]. The Appendix gives our data sources.

To investigate the behaviour of returns we form portfolios of oil stocks. We proxy the behaviour of all upstream oil stocks by the equally weighted portfolio of all stocks with return data for a given month (the "All-stocks" portfolio). The hard-to-arbitrage hypothesis suggests that sentiment-induced deviations from fundamental value should be greater in the more volatile stocks. We therefore form sub-portfolios of our sample of stocks consisting of the tercile of stocks with the highest variance of returns over the preceding 60 months and the tercile with the lowest variance. Following Baker–Wurgler, we form a portfolio that is long the tercile of high-variance stocks and short the tercile of low-variance stocks (the "Hi-Lo" portfolio). These portfolios are formed out of sample, and so represent a viable trading strategy.

We employ two measures of sentiment: the Baker–Wurgler index ("BW sentiment") and the proportion of individual investors who report that they are bullish in the regular survey conducted by the American Association of Individual Investors ("Bullish sentiment"). Both measures are available for the period July 1987 to January 2011. To facilitate comparison between the two sentiment measures, we

[1]Our results are robust to varying the definition of these variables. For example, using spot Brent prices or a closer futures contract does not affect our conclusions. Equally, we obtain qualitatively similar, but somewhat less strong, results using the Datastream index of US oil stocks rather than our equally weighted portfolio of upstream stocks only.

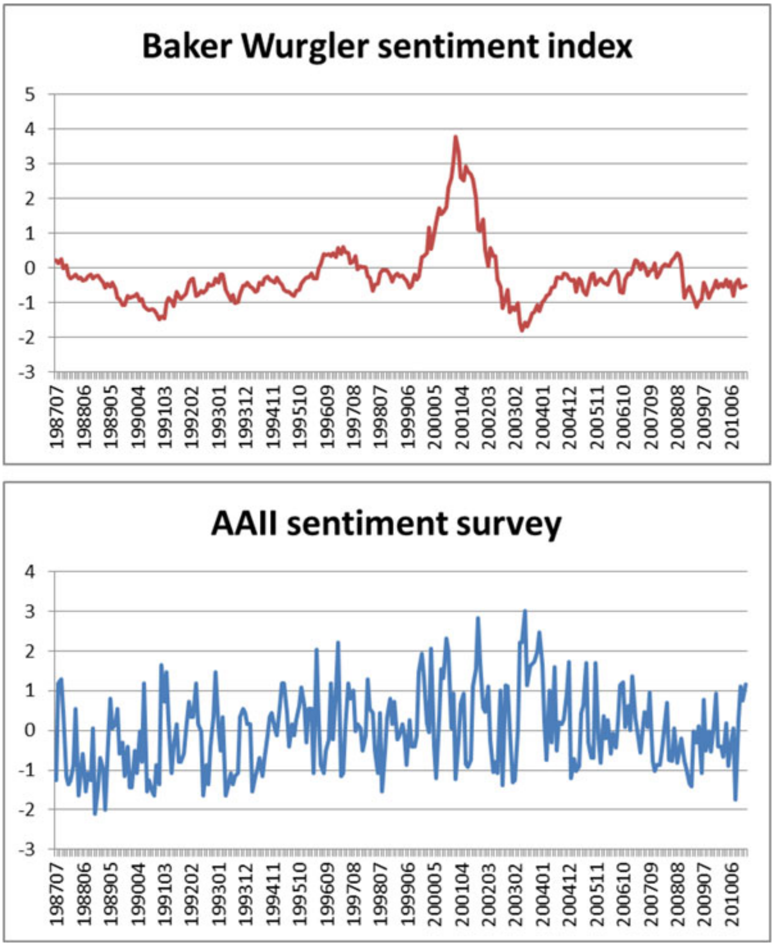


Fig. 13.1 Measures of sentiment. This figure shows plots of the Baker–Wurgler and Bullish sentiment indexes from July 1987 to January 2011. In both cases the index has been standardised to have a mean of zero and standard deviation of unity

recalibrate the index values in terms of the number of standard deviations from the mean for the total period. Figure 13.1 provides a plot of these two rescaled measures. The Baker–Wurgler index is characterised by long swings in sentiment with a marked peak in value in February 2001. By contrast, Bullish sentiment is more noisy and less persistent. The first-order autocorrelation coefficient in the Baker–Wurgler index is .96, and the serial correlation in the index persists at least through lag 6. By contrast, the first-order autocorrelation coefficient in the AAII measure is .45 and the lower order serial correlations fall away rapidly.

The monthly levels of the two sentiment indexes are only weakly related with a correlation of .09. The Baker–Wurgler index more closely resembles a cumulative

Table 13.2 Summary statistics

Panel A: Means and standard deviations of portfolio returns and changes in fundamental variables

	Lo	Hi	All	Hi-Lo	ΔWTI	ΔGas	ΔCont
Mean	.010	.014	.012	.006	.005	0.003	.0003
Std. Dev.	.069	.104	.082	.064	.093	0.196	0.036

Panel B: Correlation coefficients. Below the diagonal is the Pearson coefficient and above the diagonal in italics is the Spearman's rank coefficient. Bold face indicates coefficient is significantly different from zero at the 95 % confidence level. S_{t-1} is the Baker Wurgler sentiment index lagged 1 month, and B_{t-1} is the percentage of bullish respondents to AAII survey lagged 1 month

	Lo_t	Hi_t	All_t	Hi_t–Lo_t	S_{t-1}	B_{t-1}	ΔWTI_t	ΔGas_t	$\Delta Cont_t$
Lo_t		***.77***	***.92***	***.25***	*.02*	*.08*	***.45***	***.42***	***−.19***
Hi_t	**.82**		***.94***	***.77***	*−.04*	*.12*	***.47***	***.39***	***−.23***
All_t	**.95**	**.95**		***.54***	*−.01*	*.10*	***.48***	***.43***	***−.23***
Hi_t-Lo_t	**.30**	**.79**	**.57**		*−.08*	*.14*	***.25***	***.20***	*−.15*
S_{t-1}	.00	−.05	−.03	−.08		*.09*	*−.04*	*−.03*	*.03*
B_{t-1}	.10	**.16**	.12	.15	.08		*.14*	*.09*	*−.08*
ΔWTI_t	**.48**	**.49**	**.50**	**.30**	−.09	**.12**		***.27***	***−.71***
ΔGas_t	**.41**	**.41**	**.44**	**.24**	−.07	.11	**.24**		***−.18***
$\Delta Cont_t$	−.15	**−.21**	**−.18**	**−.21**	.06	−.07	**−.73**	−.12	

The table shows summary statistics for portfolio returns and log changes in the fundamental variables, monthly data 1988.01–2011.01. The underlying stock prices are for a balanced sample of 121 upstream oil firms. Lo is the return on the portfolio of stocks in the lowest volatility tercile. Hi is the return on the portfolio in the highest volatility tercile. All is the return on the portfolio of the entire sample. Hi–Lo is the return on the portfolio that is long the Hi portfolio and short the Lo portfolio. Δ WTI is the first difference of the log spot price of WTI. Δ Gas is the first difference of the log spot natural gas price. ΔCont is the first difference in the log of the WTI contango, where contango is defined as the ratio of the price of the sixth most distant futures contract to that of the nearest contract

sum of past values of the Bullish measure[2], which is consistent with the Baker–Wurgler index capturing cumulative deviation from fundamentals rather than short-term swings in sentiment.

Table 13.2 shows descriptive statistics for our variables. Panel A shows the means and standard deviations of the oil portfolio returns and of the changes in the fundamental values. Although the portfolios are formed out-of-sample, the standard deviation of the high-volatility portfolio is 50 % higher than that of the low-volatility portfolio. The volatility of the Hi-Lo portfolio which should, in principle, be hedged against changes in fundamentals is almost as high as that of the Hi volatility portfolio, suggesting that the long-short strategy may have only limited effect in controlling risk.

[2]A regression of the Baker–Wurgler index on the concurrent and nine lagged values of the AAII measure gives a positive loading on each of the independent variables with a multiple correlation of .34.

Panel B of Table 13.2 shows the correlation matrix for the entire period. Several features of the matrix are of interest and point to issues that are explored in more depth later.

1. The correlations between the returns on oil stocks and the three fundamental variables are quite high. Taken together, the three fundamental variables explain 41 % of the variance in the returns on the portfolio of all oil stocks.
2. The long-short portfolio of oil stocks (Hi-Lo) is not well hedged against fundamentals, and its returns remain quite highly correlated with all three fundamental variables. In other words, the high-volatility stocks are not only more difficult to arbitrage, but they also have different loadings on the fundamental factors.
3. There is little correlation between the fundamental variables and lagged sentiment. This suggests that controlling for fundamentals may not substantially change any estimate of the effect of sentiment, but also may make it easier to decompose portfolio returns into fundamental and sentiment components.
4. The correlations between returns and lagged sentiment are larger in absolute value for the high-variance portfolio and the long-short portfolio. This is consistent with the Baker–Wurgler cost-of-arbitrage hypothesis.

13.4 Sentiment and Returns

In this section we examine the influence of sentiment on returns. In Sect. 13.4.1 we provide evidence that the returns on the portfolios of oil stocks are characterised by momentum and longer-term mean reversion. We then examine the returns to see whether these patterns in returns are a function of our two measures of sentiment. In Sects. 13.4.2 and 13.4.3 we go on to decompose the returns into a fundamental and residual component and we analyse the relationship between these two components and our sentiment measures. In Sect. 13.4.4 we then examine the relationship between stock returns and "deep" fundamentals based on demand and supply in the oil market.

13.4.1 The Influence of Sentiment on the Hi-Lo Portfolio

Before testing for the effect of sentiment on returns, we first examine the serial properties of returns on the All-stock portfolio and the relationship of these returns to fundamentals. Table 13.3 shows for our portfolio of oil stocks the variance rates at differing intervals expressed as a proportion of the 1-month variance rate using the variance ratio test with overlapping data proposed by Lo and MacKinlay (1988). Consistent with standard results, the variance ratio rises for 6–9 months reflecting medium-term momentum and then falls back over the following year reflecting longer-term mean reversion.

Table 13.3 Variance ratios of the All-stocks portfolio

	Interval (months)								
	1	3	6	9	12	15	18	21	24
1983–2012	1.0	1.14	1.23	1.20	1.12	1.09	1.05	.99	.95

The table shows the variance of the All-stocks portfolio over different intervals as a proportion of the 1-month variance rate, measured over 1983.01–2012.12. The portfolio invests equal amounts each month in all US oil production and exploration stocks listed on the New York Stock Exchange

We examine the influence of sentiment by regressing total returns on the two lagged sentiment measures:

$$R_t = a + \theta_{S,R} S_{t-1+} + \theta_{B,R} B_{t-1} + e_{i,t} \tag{13.6}$$

To investigate the role of fundamentals we augment this regression with controls for the fundamental variables:

$$R_t = a + \theta_{S,R} S_{t-1+} + \theta_{B,R} B_{t-1} + \theta_{F,R} \Delta Ft + e_{i,t} \tag{13.7}$$

Where ΔFt is the vector of fundamental variables[3].

Table 13.4 reports the result of regressions (13.6) and (13.7) for the All-stock portfolio and the Hi-Lo portfolio over the period 1988–2011[4]. For the All-stock portfolio before controlling for fundamentals the coefficient on the BW sentiment measure has the predicted negative sign but is insignificant, whilst the coefficient for the Bullish measure is positive and significant, as predicted. Including the fundamentals raises the R^2 of the regression from .00 to .41. All three fundamental variables are highly significant with an unexpected 1 % increase in the oil price resulting in an increase of .59 % in the value of oil stocks. The coefficient on the change in the contango is positive suggesting that when the value of the future relative to the spot rises, there is a positive impact on the price of oil stocks. In other words, given the spot price of oil, the value of companies owning oil reserves increases when the futures price rises relative to the spot. Once these controls for fundamentals are included, the significance of the Bullish measure disappears and both sentiment measures become insignificant. The high R^2 on the regression with fundamentals and the change in the significance of the sentiment measures shows the importance of controlling for fundamentals in testing the effect of sentiment.

[3]We also estimated Eq. (13.7) using estimates of the innovations in the fundamental variables. These were estimated from an AR process with the optimal (i.e. not pre-specified) number of lags. The results were not sensitive to whether the fundamental variables were whitened.

[4]Note that the regressions employ data only from 1988. The first 60 months of data are used to form the initial high- and low-variance portfolios.

Table 13.4 Regression of portfolio returns on lagged sentiment with and without controls for fundamentals

	All-stocks portfolio		Hi-Lo portfolio	
	No controls	Fundamental controls	No controls	Fundamental controls
Baker–Wurgler sentiment	−.003 (−.69)	.002 (.51)	−.006 (−1.47)	−.003 (−.91)
Bullish sentiment	.010 (2.11)**	.003 (.70)	.010 (2.68)***	.007 (1.91)*
Oil return		.587 (9.81)***		.170 (2.97)***
Gas return		.129 (6.54)***		.053 (2.82)***
Oil contango change		.805 (5.35)***		.018(.13)
Rbar^2	.01	.41	.02	.12

The table summarises the results of a regression of returns on each of two oil portfolios against two lagged sentiment measures, S_{t-1} and B_{t-1}, for the period 1988.01–2011.01. The coefficients on the sentiment measures are shown without and with controls for the effect of fundamentals. The All-stocks portfolio invests equal amounts each month in all US oil production and exploration stocks listed on the New York Stock Exchange. The Hi-Lo portfolio has long positions in the tercile of stocks with the highest variance over the prior 60-month period and short positions in the tercile of stocks with the lowest variance over the prior 60-month period. The regressions are estimated from monthly data using OLS with a constant (not reported). t-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance

The hard-to-arbitrage hypothesis predicts that the influence of sentiment will be most marked for high-variance stocks. Therefore, following Baker–Wurgler, we also report in Table 13.4 the result of regressions (13.6) and (13.7) for the portfolio that is long high-variance stocks and short low-variance stocks. The high-variance stocks are characterised by a heavier loading on the three fundamental factors. This may stem from the higher relative importance of exploration (as opposed to production) activity as proxied by a low ratio of revenues to assets.[5] Thus, despite its apparent hedged position, the Hi-Lo portfolio remains significantly exposed to oil and gas returns After controlling for fundamentals, the coefficients on both sentiment measures have the predicted sign, but only the Bullish variable is significant at the 10 % level.

For the Hi-Lo portfolio we also estimate these regressions for two sub-periods divided at 2000. The motivation for looking separately at the sub-periods is the sharp increase in institutional investment in oil futures after 2000 (Buyuksahin and Robe (2012) and Singleton (2011)). In the ten years to 2010 open interest in crude oil futures by non-commercial traders was 5.6 times its level over the previous 15 years (Fig. 13.2). Figure 13.3 shows that there was also a sharp rise in the cumulative cash flows into managed futures funds, which increased from $9.3 billion in September 2002 to $137.0 billion in March 2008 before losing most of these gains in 2009.

[5]It is also possible that the lower exposure of low-variance stocks to energy prices reflects hedging activity, although Haushalter (2000) suggests that hedging is more commonly used by the more risky oil and gas firms.

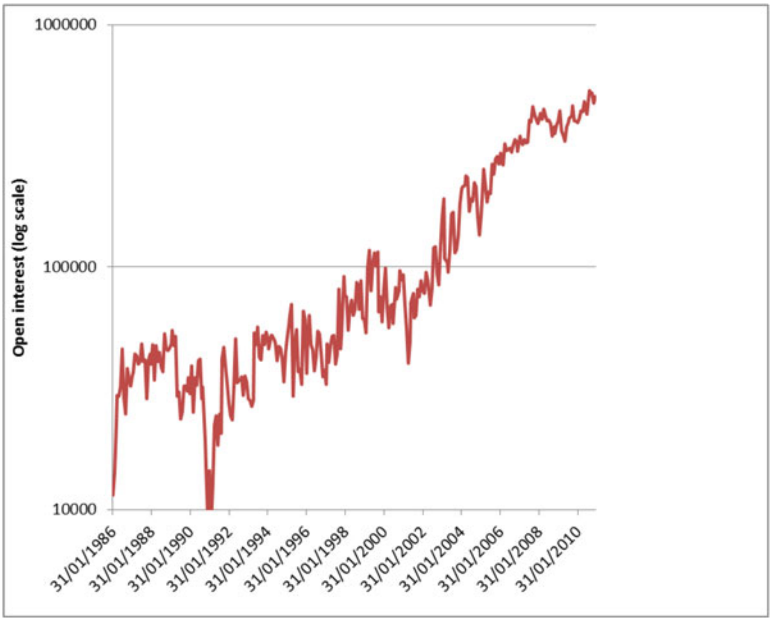


Fig. 13.2 Open interest in oil futures. This figure shows on a log scale the level of open interest in crude oil futures by non-commercial traders from January 1986 to December 2012. Source: CFTC

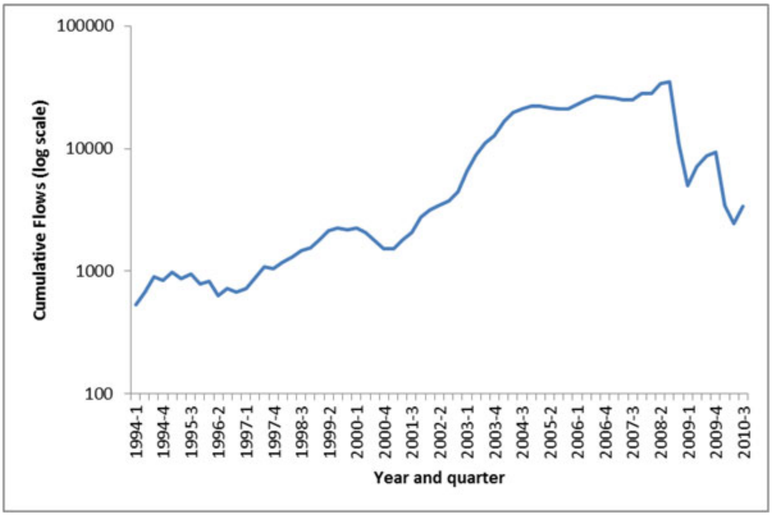


Fig. 13.3 Cumulative flows into commodity hedge funds. This figure shows on a log scale the cumulative quarterly cash flows into Managed Futures Funds in millions of dollars from first quarter 1994 to the third quarter 2010. Source: TASS Research

Table 13.5 Sub-period regressions of Hi-Lo portfolio returns on lagged sentiment with and without controls for fundamentals

	Hi-Lo portfolio	
	No controls	Fundamental controls
	Panel A: 1988–2000	
Baker–Wurgler sentiment	.008 (1.06)	.007 (.85)
Bullish sentiment	.004 (.63)	.001 (.17)
Oil return		.214 (2.23)**
Gas return		.093 (2.59)**
Oil contango change		.085 (.40)
Rbar^2	.00	.10
	Panel B: 2001–2011	
Baker–Wurgler sentiment	−.009 (−.2.41)**	−.007 (−1.90)*
Bullish sentiment	.014 (2.76)***	.011 (2.28)**
Oil return		.150 (2.28)**
Gas return		.023 (1.28)
Oil contango change		−.016 (−.08)
Rbar^2	.10	.18

The table summarises the results of a regression of returns on the Hi-Lo portfolio against two lagged sentiment measures, S_{t-1} and B_{t-1}, for the two sub-periods (1988.01–2000.12 and 2001.01–2011.01). The coefficients on the sentiment measures are shown without and with controls for the effect of fundamentals. The Hi-Lo portfolio invests in US oil production and exploration stocks listed on the New York Stock Exchange. It holds long positions in the tercile of stocks with the highest variance over the prior 60-month period and short positions in the tercile of stocks with the lowest variance over the prior 60-month period. The regressions are estimated from monthly data using OLS with a constant (not reported). *t*-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance

Table 13.5 repeats regressions (13.6) and (13.7) for the two sub-periods. There is a substantial difference between the sub-periods both in the coefficients and their significance. In the earlier period even before controlling for fundamentals neither of the sentiment coefficients is significant and the coefficient for the Baker–Wurgler measure has the wrong sign. In the second period, even after controlling for fundamentals, both sentiment measures have the predicted signs and remain significant at the 10 % level or better.

13.4.2 Tests Using Fundamentals and Deviations from Fundamentals

We now test the hypothesis that sentiment operates by affecting deviations from fundamental value. To do so, we first decompose returns on the Hi-Lo portfolio into a fundamental component and a residual, and we then estimate the relationship between our sentiment measures and each of these two components. Since the loadings of this portfolio on the fundamental factors may be time-varying we estimate the loadings using rolling 60-month regressions of portfolio returns on the fundamental variables:

$$R_{it} = a_i + b_i \Delta \text{ WTI}_t + c_i \Delta \text{ Gas}_t + d_i \Delta \text{ Cont}_t + u_{it} \qquad (13.8)$$

We use the coefficients from this regression over the prior 60 months combined with the change in the month's fundamentals to estimate that part of the month's return that was due to fundamentals. The difference between the actual return in a month and the fundamental return is the residual, or non-fundamental, return.

We estimate the VAR system Eqs.(13.2)–(13.5) using GMM with the Newey–West correction for standard errors. Table 13.6 shows the results for the entire period 1988–2011[6]. Contrary to the "deviations from fundamentals" hypothesis all the coefficients of the regression of these deviations on lagged variables in column (2) are insignificant, and the Rbar^2 is negative. In contrast, the regression of fundamental returns on lagged sentiment (column (1)) has an Rbar^2 of 5 %. There is a negative coefficient on lagged B–W sentiment and a positive coefficient on the lagged bullish variable. Both coefficients are strongly significant. This result is consistent with sentiment-based trading operating largely through the fundamentals

Table 13.6 VAR of fundamental values for the Hi-Lo portfolio, deviations from fundamentals, Baker–Wurgler sentiment, and Bullish sentiment

	(1)	(2)	(3)	(4)
	Dependent variable			
	$F_t - F_{t-1}$	$NF_t - NF_{t-1}$	S_t	B_t
$F_{t-1} - F_{t-2}$	−.0312 (−.48)	.0349 (.27)	2.1492 (3.09)***	−1.0366 (−.43)
$NF_{t-1} - NF_{t-2}$	−.0014 (−.06)	.0109 (.19)	.3230 (1.17)	−.8704 (−.93)
S_{t-1}	−.0045(−2.43)**	−.0010 (−.25)	.9668 (33.78)***	.0543 (.75)
B_{t-1}	.0047 (3.17)***	.0054 (1.49)	.0240 (1.50)	.4478 (6.68)***
Rbar^2	.05	−.01	0.92	0.20

The table shows a VAR of the return to fundamentals ($F_t - F_{t-1}$) for the Hi-Lo portfolio, return to deviations from fundamentals ($NF_t - NF_{t-1}$) for the Hi-Lo portfolio, Baker–Wurgler sentiment (S_{t-1}), and Bullish sentiment (B_{t-1}). Monthly data 1988.01–2011.01. The Hi-Lo portfolio has long positions in the tercile of stocks with the highest variance over the prior 60-month period and short positions in the tercile of stocks with the lowest variance over the prior 60-month period. Returns on the portfolio each month are split between fundamental returns and deviations from fundamentals. Fundamental returns are based on a regression of the portfolio return on the oil price return, gas price return, and the change in contango in the oil market, using a rolling window of 60 months prior to the month for which the stock return is split. The system is estimated using GMM with Newey–West correction for the standard errors. *t*-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance

[6]The period is reduced by 5 years because the first 60 months are used to estimate Eq. (13.8).

themselves, rather than through the deviations in the share price from fundamentals. The negative coefficient on lagged B-W sentiment is consistent with high sentiment signalling that the oil market is above its equilibrium and likely to fall. The positive coefficient on the Bullish variable is consistent with short-term momentum pushing the oil market upwards when retail investors are bullish.

Column (3) shows that the Baker–Wurgler sentiment variable is persistent, with a partial serial correlation of 0.97 consistent with a half-life of 20 months. The variable also responds to lagged changes in fundamentals, but not to lagged changes in deviations from fundamentals. Again, this is consistent with sentiment operating through the fundamentals themselves and not through deviations of share prices from fundamentals. Column (4) shows that the Bullish variable is much less persistent, with a half-life of less than a month. It has positive serial correlation and positively responds to past fundamental returns.

Table 13.7 decomposes the Hi-Lo portfolio data into two sub-periods, divided at the end of 2000. Column (1) shows the results for the entire period, and columns (2) and (3) for the two sub-periods. The two sentiment variables have no significant effect on deviations from fundamentals in any period, though the coefficients on the Bullish variable consistently have the correct sign. The coefficients from the regression of the fundamental component of returns on the two lagged sentiment variables have the correct sign but are insignificant in the first sub-period. By

Table 13.7 Sub-period VAR results

	(1)	(2)	(3)
	1988–2011	1988–2000	2001–2011
Regression of fundamental return on lagged sentiment			
Baker–Wurgler sentiment: $\theta_{S,F}$	−.0045 (−2.43)**	−.0011 (−.41)	−.0059 (−2.52)**
Bullish sentiment: $\theta_{B,F}$	.0047 (3.17)***	.0022 (1.21)	.0069 (2.60)***
Rbar^2	.05	−.01	.08
Regression of deviations from fundamentals on lagged sentiment			
Baker–Wurgler sentiment: $\theta_{S,\mathrm{NF}}$	−.0010 (−.25)	.0092 (1.29)	−.0046 (−1.06)
Bullish sentiment: $\theta_{B,\mathrm{NF}}$	.0054 (1.49)	.0018 (.33)	.0049 (1.20)
Rbar^2	−.01	−.01	.01
Partial serial correlation coefficient of			
Baker–Wurgler sentiment: $\theta_{\mathrm{S,S}}$	.9668 (33.78)***	1.008 (31.31)***	.9555 (30.60)***
Bullish sentiment: $\theta_{\mathrm{B,B}}$	.4478 (6.68)***	.3583 (5.65)***	.4383 (4.02)***

The table shows a VAR of the return to fundamentals ($F_t - F_{t-1}$) for the Hi-Lo portfolio, return to deviations from fundamentals ($NF_t - NF_{t-1}$) for the Hi-Lo portfolio, Baker–Wurgler sentiment, S, and bullish sentiment, B. Monthly data 1988.01–2011.01, divided at 2000.12. Stock returns each month are split between fundamental returns and deviations from fundamentals based on a regression of the stock return on the oil price return, gas price return, and the change in contango in the oil market, using a rolling window of 60 months prior to the month for which the stock return is split. Coefficient $\theta_{i,j}$ measures the autoregression coefficient of variable j on the lagged value of variable *i*. The system is estimated using GMM with Newey–West correction for the standard errors. *t*-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance contrast, in the second period the corresponding coefficients are strongly significant. Although the time-series behaviour of the sentiment variables appears to be the similar in the two sub-periods, the effect of sentiment on stock returns changes completely in the second period. Consistent with the result that sentiment affects prices largely via fundamentals, the effect occurs only once there is significant investment interest in the fundamental markets post-2000[7].

13.4.3 The Effect of the Differencing Interval

Table 13.8 shows the effect of increasing the differencing interval to 3 months and 12 months. The results are shown for the entire period (Panel A) and the two sub-periods (Panels B and C). The VAR is estimated using GMM with overlapping observations and Newey–West corrected standard errors. In the regression of fundamental returns the effect of moving from a 1-month to 3-month differencing interval is to increase the magnitude and significance on both of the lagged sentiment measures in all three periods. The Rbar^2 of the regression of fundamentals on sentiment increases dramatically, rising to .29 for a 12-month differencing interval in the second sub-period. Thus the sentiment measures appear to have a prolonged effect on the fundamental returns. In contrast, the longer differencing intervals have almost no effect on the coefficients for the deviations from fundamentals, which remain insignificant at all intervals and in all periods.

13.4.4 Deep Fundamentals

Our measure of the fundamental return on the portfolio of oil stocks is equivalent to a weighted average of the contemporaneous change in the price of oil and gas and the change in the contango. The evidence that this weighted average is a function of the prior level of sentiment implies that oil and gas prices are themselves influenced by sentiment (Pindyck, 1993). Thus it appears that sentiment drives oil prices away from equilibrium values in a way that leads to predictable returns on oil stocks. This effect increases after 2000 when interest in commodities as an asset class increased significantly. 1-month returns are slightly predictable using sentiment, but returns over a 1-year horizon are highly predictable. Overall, the results appear to reflect a slowly changing but predictable component of oil and gas prices that is related to

[7]The sharp changes in cumulative flows into commodity hedge funds prompted us to examine (more in hope than expectation) the effect within the VAR of interacting the cumulative flows with the sentiment variables. There was no evidence that the impact of sentiment on returns was related to the cumulative flows into managed futures funds.

Table 13.8 Effect of sentiment on returns using different measurement intervals

Differencing interval	Fundamentals			Deviations from fundamentals		
(Months)	$\theta_{S,F}$	$\theta_{B,F}$	Rbar^2	$\theta_{S,\text{NF}}$	$\theta_{B,\text{NF}}$	Rbar^2
Panel A: 1988–2011						
1	−.0045 (−2.43)**	.0047 (3.17)***	.05	−.0010 (−.25)	.0054 (1.49)	−.01
3	−.0103 (−2.61)***	.0116 (3.64)***	.15	−.0033 (−.26)	.0064 (.94)	−.01
12	−.0482 (−3.12)***	.0271 (2.79)***	.23	−.0212 (−.82)	−.0110 (−.60)	.02
Panel B: 1988–2000						
1	−.0011 (−.41)	.0022 (1.21)	−.01	.0092 (1.29)	.0018 (.33)	−.01
3	−.0035 (−.61)	.0081 (2.41)**	.14	.0241 (1.21)	−.0013 (−.11)	.02
12	−.0449 (−1.76)	.0157 (1.83)	.16	−.0465 (−.98)	.0286 (1.15)	.08
Panel C: 2001–2011						
1	−.007 (−2.52)**	.0069 (2.60)***	.08	−.0046 (−1.06)	.0049 (1.20)	.01
3	−.0154 (−3.54)***	.0163 (3.10)***	.27	−.0074 (−.57)	.0061 (.59)	−.00
12	−.0648 (−3.12)***	.0241 (1.51)	.29	−.0034 (−.11)	−.0199 (−.76)	.06

The table shows results from a VAR of the return to fundamentals ($F_t - F_{t-1}$) for the Hi-Lo portfolio, return to deviations from fundamentals ($NF_t - NF_{t-1}$) for the Hi-Lo portfolio, Baker–Wurgler sentiment (S), and Bullish sentiment (B). The differencing interval in months is shown in the left column. Monthly data 1988.-01 to 2011.01, divided at 2000.12. Stock returns each month are split between fundamental returns and deviations from fundamentals based on a regression of the stock return on the oil price return, gas price return, and the change in contango in the oil market, using a rolling window of 60 months prior to the month for which the stock return is split. The system is estimated using GMM with Newey–West correction for the standard errors. *t*-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance sentiment and generates a predictable return on oil stocks. Sentiment appears to have no effect on the price of oil stocks other than through the prices of the commodities themselves.

We can gain some further insight by examining the relationship between the change in oil prices and prior sentiment while controlling for the deeper fundamentals that determine oil prices.

$$\Delta \text{WTI}_t = a + \theta_S S_{t-1} + \theta_B B_{t-1} + \theta_F \Delta DF_t + \text{e}_t \tag{13.9}$$

where ΔDF_t is the vector of the underlying determinants of the change in oil prices.

The main problem in estimating Eq. (13.9) is the lack of good proxies for deeper fundamentals that are available at sufficiently high frequency. We proxy the fundamental determinants oil prices by changes in world oil production and consumption, changes in world proven reserves (annual data only), changes in oil inventories (monthly data only), and a measure of economic growth. We estimate Eq. (13.9) using annual data, overlapping 12-month data and overlapping 3-month data. In the case of the annual data estimates are for the period 1988–2011 and in the case of the regressions using monthly data estimates are for the period 1994–2011. The results are summarised in Table 13.9.

The controls for fundamental variables have little explanatory power in the overlapping 3-month data regressions but in each case the coefficients on the Baker–Wurgler measure are negative and those on the Bullish measure are positive. However, with relatively few independent observations, the tests lack power and in only two cases is the coefficient on sentiment significant at the 10 % level. Thus the table provides mild direct support that sentiment affects energy prices and thereby the return on oil stocks.

13.5 Robustness Tests

We have already noted that our findings are robust to (a) pre-whitening the fundamental variables, (b) using different definitions of the crude oil price and the oil contango, (c) using the Datastream index of oil stocks.

13.5.1 Long-Only Portfolios

To the extent that the Hi-Lo portfolio is better hedged against fundamental factors than long-only portfolios, the fundamental component of returns will be relatively small. We therefore repeated the VAR estimates with long-only portfolios. The results for the tercile of stocks with the highest variance were very similar to those for the Hi-Lo portfolio. In particular, the effects of sentiment on returns were significant only for the second period, and sentiment impacted returns largely through the fundamental component.

Table 13.9 Regression of changes in the oil price on "deep" fundamental factors and lagged sentiment

	Annual data (1988–2011) $N = 24$	Overlapping 12-month data (1995.01–2011.01) $N = 193$	Overlapping 3-month data (1994.04–2011.01) $N = 202$
Baker–Wurgler sentiment	−.16 (-1.19)	−.05 (−1.52)	−.03 (−1.90)*
Bullish sentiment	.99 (1.68)*	.03 (1.32)	.01 (.74)
Δ world oil prod'n	−14.78 (−2.60)**	.80 (.68)	.21 (.17)
Δ oil consumption[a]	24.88 (2.62)**	−1.20 (−.1.12)	−1.01 (−2.20)**
Δ economic activity[b]	−6.56 (−.84)	2.02 (2.33)**	3.62 (1.28)
Δ world proven oil reserves	.92 (.38)		
Δ OECD oil stocks		−6.33 (−5.33)***	−2.12 (−3.48)***
$Rbar^2$	.40	.49	.13

The table summarises the results of a regression of changes in the price of WTI against two lagged sentiment measures, S_{t-1} and B_{t-1}, and controls for fundamental factors. In the annual regressions the dependent variable consists of 31 annual changes in the oil price; in the overlapping 12-month regressions they consist of 193 rolling 12-month changes. The regressions are estimated using OLS with a constant (not reported). Standard errors in the overlapping regressions are calculated using a Newey–West correction. *t*-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance.

[a]World consumption is used in the annual regression, OECD consumption in the overlapping regression

[b]World GDP is used in the annual regression, US industrial production in the overlapping regression

13.5.2 Nasadaq Stocks

To test further the robustness of these results to a different sample of oil stocks, we extended the sample to include 274 stocks of US oil and exploration companies that were traded on Nasdaq. Although this produced a larger sample, the quality of the Nasdaq data appears to be inferior with shorter time series for many stocks, leptokurtic returns and more zero returns. Unsurprisingly, the high variance portfolio tended to have a high concentration in Nasdaq stocks.

Table 13.10 summarises estimates of regression Eq. (13.8) for the expanded sample of NYSE and Nasdaq stocks. The expanded portfolio is better hedged against fundamental factors and the addition of these factors has therefore less effect on estimates of the sentiment effect. Otherwise, the results are similar to those reported in Tables 13.4 and 13.5. The coefficients on *B–W* are consistently negative and those on Bullish consistently positive. However, there continues to be a big difference between the two periods with the coefficients being significant only in the later period.

13.5.3 The Effect of Lagged Market Returns

To evaluate the role of market returns in generating sentiment, we augmented the VAR system by including the lagged market return in each of the regressions. This did not significantly change the relationships between either of the sentiment variables and either of the returns. It did not increase the R^2's for the prediction of returns. The Bullish sentiment variable is not significantly related to the lagged market return in the second sub-period, where the main sentiment effect is apparent. This suggests that the momentum generated by the positive relationship between fundamental returns and lagged Bullish sentiment is not simply a proxy for the effect of lagged market returns.

13.5.4 Lagged Fundamentals

We also added to the VAR system more lags of the fundamental returns. These were generally insignificant and did not change the basic results. The sentiment variables remained significant in the second sub-period and the effect of sentiment showed up only in the fundamental regression and not the deviations from fundamentals.

Table 13.10 Regression of portfolio returns on lagged measures of sentiment with and without controls for fundamentals. The portfolio is an expanded long-short portfolio of NYSE and Nasdaq oil stocks (the Hi-Lo portfolio)

	Hi-Lo portfolio	
	No controls	Fundamental controls
	Panel A: 1988–2011	
Baker–Wurgler sentiment	−.012 (−3.69)***	−.012 (−3.57)***
Bullish sentiment	.007 (2.14)**	.007 (2.00)**
Oil return		.101 (1.93)*
Gas return		−.018 (−1.04)
Oil contango change		.130 (.99)
$Rbar^2$	.05	.06
	Panel B: 1988–2000	
Baker–Wurgler sentiment	−.008 (1.20)	−.006 (−.90)
Bullish sentiment	.003 (.60)	.003 (.64)
Oil return		.042 (.51)
Gas return		−.046 (−1.50)
Oil contango change		.032 (.17)
$Rbar^2$	−.00	−.01
	Panel C: 2001–2011	
Baker–Wurgler sentiment	−.013 (−.3.23)***	−.012 (−3.01)***
Bullish sentiment	.011 (2.24)**	.010 (2.04)**
Oil return		.140 (2.08)**
Gas return		−.010 (−.49)
Oil contango change		.178 (86)
$Rbar^2$	.11	.13

The table summarises the results of a regression of returns on a portfolio of oil stocks against two lagged sentiment measures, S_{t-1} and B_{t-1}, for the period 2001.01–2011.01. The portfolio is constructed from an expanded sample of 395 upstream oil stocks traded on the NYSE and Nasdaq. The coefficients on the sentiment measures are shown without and with controls for the fundamentals. The Hi-Lo portfolio has long positions in the tercile of stocks with the highest variance over the prior 60-month period and short positions in the tercile of stocks with the lowest variance over the prior 60-month period. The regressions are estimated from monthly data using OLS with a constant (not reported). *t*-statistics in parentheses. * denotes 10 % significance, ** 5 % significance, and *** 1 % significance

13.6 Conclusions

Using a sample of upstream oil stocks where we have a good proxy for fundamental value, we show that sentiment predicts returns. However, the effect is highly time-varying, appearing only after the post-2000 increased interest in oil-related assets.

Sentiment effects come in two forms: retail investor sentiment predicts short-term momentum, and Baker–Wurgler sentiment predicts medium-term mean reversion of fundamental factors. Whilst the sentiment variables explain only a negligible proportion of the variance of returns, the additional return due to a change in sentiment is not unimportant. For example, in the second period for the portfolio of all oil stocks a one standard deviation rise in the level of investor sentiment added about .3 % to the following month's return; during the same period a similar one-standard-deviation rise in the Baker–Wurgler index *reduced* return by about .3 %.

Contrary to the hard-to-arbitrage hypothesis, sentiment affects returns on these stocks through fundamentals rather than through deviations from fundamentals. Overall, it appears that retail sentiment drives the prices of oil and gas futures away from their deeper fundamental values until the deviation is sufficiently large that arbitrageurs drive the prices back towards their equilibrium values. This process for the fundamentals is then reflected in the prices of upstream oil stocks.

These effects appear even in a portfolio that is long hard-to-arbitrage stocks and short easy-to-arbitrage stocks, because this portfolio has a net exposure to fundamentals. This has implications for tests of the hard-to-arbitrage hypothesis, showing that it is important to have effective controls for fundamentals even when the long-short portfolio is used.

Our finding that sentiment affects upstream oil stocks through the fundamentals raises the issue as to whether this is also the case with other industries that invest in assets that are traded in speculative markets. Obvious examples would be stocks in other extractive industries but a similar effect could characterise financial institutions. It also prompts the question whether the magnitude of any sentiment effects depends on the extent to which the fundamentals are tradeable. If this is the case, sentiment effects might vary not just with ease of arbitrage but with the nature of the company's fundamentals. The sharp increase in the significance of sentiment effects in the post-2000 period was accompanied by an increase in speculative activity in energy futures. If these effects are truly linked, then it raises the question as to the effect of trading activity on the influence of sentiment. These would appear to be fruitful, if difficult, areas for future research.

Acknowledgements The authors appreciate the comments of Bernell Stone, the reviewer of the manuscript.

Appendix: Principal Data Sources

Stock samples: All common stocks with NAIC code of 211111 or SIC code of 1311 (oil production or exploration) that were listed on the NYSE or Nasdaq and whose issuers were incorporated in the USA. Returns data were taken from the CRSP monthly database. Portfolio returns were constructed from equally weighted holdings in all stocks with valid returns data for that month. Portfolio returns were then converted to continuously compounded returns.

Oil prices: Month-end spot prices for West Texas Intermediate taken from the Energy Administration website at www.eia.gov.

Natural gas prices: Monthly spot prices for Natural Gas Wellhead Price taken from Globalfindata. Prices are month averages from March 1983–December 1995 and end-of-month from January 1996–January 2011.

Contango: NYMEX futures prices are taken from Quandl at www.quandl.com. The contango is defined as the price of the contract that is sixth nearest to delivery divided by the price of the contract that is closest to delivery. The change in contango is defined as $\ln(\text{contango}_t) - \ln(\text{contango}_{t-1})$.

Baker–Wurgler Sentiment Index: SENT[1] constructed from IPO volume, IPO first-day returns, market turnover, and the market-book ratio of high-volatility stocks relative to that of low-volatility stocks. See http://people.stern.nyu.edu/jwurgler/. The index is rescaled to have mean zero and unit standard deviation.

American Association of Individual Investors (AAII) Investor Sentiment Survey: Proportion of investors reporting they are bullish divided by the total proportion reporting that they are either bullish or bearish (i.e. not neutral). Taken from final week's survey in each month as reported on www.aaii.com/sentimentsurvey. The index is rescaled to have mean zero and unit standard deviation. Data are available from July 1987.

References

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance, 68*, 929–981.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance, 61*, 1645–1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives, 21*, 129–151.

Baker, M., & Wurgler, J. (2012). Comovement and predictability relationships between bonds and the cross-section of stocks. *Review of Asset Pricing Studies, 1*, 57–87.

Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics, 104*, 272–287.

Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics, 75*, 283–317.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*, 307–343.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance, 11*, 1–27.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business, 78*, 405–440.

Brown, S. J., Goetzmann W. N., Hiraki, T., Shirishi, N. & Watanabe, M. (2003). *Investor sentiment in U.S. and Japanese daily mutual fund flows*. NBER working paper No. W9470.

Buyuksahin, S. & Robe, M.A. (2012). Speculators, commodities and cross-market linkages. Retrieved from SSRN: http://ssrn.com/abstract = 1707103 or http://dx.doi.org/10.2139/ssrn.1707103.

Chen, N.-f., Kan, R., & Miller, M. (1993). Are the discounts on closed end funds a sentiment index? *Journal of Finance, 48*, 795–800.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1991). Speculative dynamics. *Review of Economic Studies, 58*, 529–546.

Daniel, K. D., Hirshleifer, D. A., & Subrahmanyam, A. (1998). A theory of overconfidence, self-attribution, and security market under- and over-reactions. *Journal of Finance, 53*, 1839–1885.

Derrien, F., & Kecskés, A. (2009). How much does investor sentiment really matter for equity issuance activity? *European Financial Management, 15*, 787–813.
Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy, 96*(April), 246–273.
Haushalter, G. D. (2000). Financing policy, basis risk, and corporate hedging. *Journal of Finance, 55*, 107–152.
Jegadeesh, N. (1990). Evidence of predictable behavior of stock returns. *Journal of Finance, 45*, 881–898.
Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance, 48*, 65–91.
Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance, 56*, 699–720.
Kumar, A., & Lee, C. M. C. (2006). Retail investor sentiment and return comovements. *Journal of Finance, 61*, 2451–2486.
Lee, C. M. C., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance, 46*, 75–109.
Lehmann, B. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics, 105*, 1–28.
LeRoy, S. F., & Porter, R. D. (1981). The present value relation: tests based on implied variance bounds. *Econometrica, 49*, 97–113.
Lo, A. W., & MacKinlay, C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies, 1*, 41–66.
Miller, M. H., & Upton, C. W. (1985a). A test of the Hotelling valuation principle. *Journal of Political Economy, 93*(February), 1–25.
Miller, M. H., & Upton, C. W. (1985b). The pricing of oil and gas: Some further results. *Journal of Finance, 40*, 1009–1018.
Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2011). Time series momentum. *Journal of Financial Economics, 104*, 228–250.
Neal, R., & Wheatley, S. M. (1998). Do measures of sentiment predict returns? *Journal of Financial and Quantitative Analysis, 33*, 523–535.
Pindyck, R. S. (1993). The present value model of rational commodity pricing. *The Economic Journal, 103*, 511–530.
Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics, 22*, 27–60.
Qiu, L. & Welch, I. (2004). *Investor sentiment measures.* NBER Working Paper No. W10794.
Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review, 71*, 421–436.
Singleton, K. J. (2011). *Investor flows and the 2008 boom/bust in oil prices*. Retrieved from SSRN: http://ssrn.com/abstract = 1793449 or http://dx.doi.org/10.2139/ssrn.1793449.
Swaminathan, B. (1996). Time varying expected small firm returns and closed-end fund discounts. *The Review of Financial Studies, 1*, 845–887.

Chapter 14
Constructing Mean Variance Efficient Frontiers Using Foreign Large Blend Mutual Funds

Ganlin Xu, Harry Markowitz, Minyee Wang, and John B. Guerard

14.1 Introduction

Mutual funds are efficient investment vehicle for small investors. The buy and hold strategy is the prevailing way of investing in mutual funds because of the trading cost and tax considerations. Since the emergence of online trading platforms, the trading cost has come down significantly. Now it is the time to evaluate strategies of more actively managed portfolios of mutual funds. In this study, we show how to use mean-variance portfolio selection methods to construct and manage portfolios of mutual funds, with the focus on funds categorized as foreign large blend by Morningstar. There are two reasons we choose this category of mutual funds. First, total foreign equity markets are as large as the US equity market now, and mutual funds are still the best way to get exposures to it. Second, this category of mutual fund is under-studied. Most researchers focus on the relative performance of US equity mutual funds. We report that: (1) The performance predictive variables that work for US equity mutual funds can also work for foreign large blend mutual funds; (2) the mean-variance approach can effectively diversify the risk of portfolios for this category of mutual funds too. The risk of the minimum variance portfolio could be 6 percentage points less than the risk of the expected-return maximizing portfolio while the realized return is only about 2 percentage points less; (3) the mean-variance approach can produce portfolios with higher Sharpe ratios than the Sharpe ratio of either the index funds or the category average which are the benchmark of

G. Xu (✉) • H. Markowitz • M. Wang
GuidedChoice.com, 8910 University Center Lane, Suite 400, San Diego, CA 92122, USA
e-mail: gxu1@guidedchoice.com

J.B. Guerard
McKinley Capital Management, LLC, 3301 C Street, Suite 500, Anchorage, AK 99503, USA

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_14

this study. Some efficient portfolios can outperform these two benchmarks by more than 2 percentage points while having the same risk levels even after transaction cost.

This paper is organized as following. In Sect. 14.2, we briefly state the single-period mean-variance portfolio selection problem with turnover constraints, and describe how we compute the *ex post* mean-variance efficient sets. In Sect. 14.3, we present two variance-covariance models and also a broad review of expected return models used for US equity mutual funds. In Sect. 14.4, we describe the data source, define the investable universe, and discuss the assumptions on transaction cost. In Sect. 14.5, we present various *ex post* efficient frontiers by varying expected return models, risk models, turnover constraints, and upper bounds. We conclude that various expected return models can be used as input for mean-variance optimizations to generate 2 percentage points more returns than benchmarks while having the same or less risk.

14.2 Single-Period and *Ex Post* Mean Variance Efficient Frontier

The mean variance portfolio construction method proposed by Markowitz (1952, 1959) assumes that an investor should maximize the expected portfolio return for a given risk level, or equivalently, minimize the risk for a given expected portfolio return. The complete efficient portfolio set can be traced out by the quadratic problem:

$$\min x^T C x - \lambda_E \mu^T x \tag{14.1}$$

where μ is the expected return vector, C is the variance-covariance matrix, and x is the portfolio weights, $\lambda_E \geq 0$ is the risk return trade off parameter. The choice of λ_E reflects the investor's risk tolerance. The more risk adverse an investor is, the smaller would be his λ_E.

In this study, we impose the no-shorting constraint, i.e., $x_i \geq 0$ for all i, an universal upper bound, i.e.,

$$x_i \leq \text{ upper bound}, \quad \text{for all } i, \tag{14.2}$$

and budget constraints $\sum x_i = 1$, i.e., fully invested. We also consider turnover constraints in terms of total buying and selling

$$\left(\sum\nolimits_{i=1}^{N} \left|x_i - x_i^o\right|\right) / 2 \leq \text{ turnover} \tag{14.3}$$

where x_i^o is previous period's portfolio weight of security i in the universe.

Using the mean-variance framework requires the investor to estimate the mean-variance parameters (μ, C), and make decisions on upper bounds, turnover, and λ_E. We refer the estimation methodology of (μ, C) together with decisions on upper bounds and turnover constraints as an investment strategy. For each period t and λ_E, the transaction cost-adjusted return $R(t, \lambda_E)$ can be easily calculated for the corresponding efficient portfolio generated at the beginning of the period. The λ_E parametric realized mean-variance set of returns $R(t, \lambda_E)$ (for $t = 0$ to period T) generated by the same strategy for all λ_E are called the *ex post* efficient frontier. We will use *ex post* efficient frontier as the criteria to evaluate investment strategies.

14.3 Risk Models and Expected Return Models

The multifactor linear model is the standard risk model. Here we assumes that the return of any mutual fund can be modeled by Eq. (14.4) with four factors, i.e., market factor R_{M}, size factor R_{SMB} (small cap portfolio minus large cap portfolio), value factor R_{HML} (high Book/Price portfolio minus low Book/Price portfolio), and momentum factor R_{WML} (last year winner portfolio minus loser portfolio).

$$R_p - R_f = \alpha_p + \beta_{\mathrm{mp}}\left(R_M - R_f\right) + \beta_{\mathrm{smbp}} * R_{\mathrm{SMB}} + \beta_{\mathrm{hmlp}} R_{\mathrm{HML}} + \beta_{\mathrm{wmlp}} * R_{\mathrm{WML}} + \epsilon_p \tag{14.4}$$

Once the betas were estimated by regressing mutual fund returns to factor returns, the variance-covariance risk matrix C can be calculated by

$$C = \beta F \beta' + \sum \tag{14.5}$$

where F is variance-covariance of factor returns and $\sum$ is the residual diagonal risk matrix. Fama and French (1992, 1996) developed this factor model for stocks. It has been adopted by mutual fund researchers since Carhart (1997). The factor returns used in this study are downloaded from French's data library. Table 14.1 shows the sample statistics of the factor returns.

The relatively low correlations and negative correlations among the factor returns make it a very attractive risk model. One can skip the regressions in Eq. (14.4) and calculate the variance-covariance matrix directly from fund's historical returns by Eq. 14.6.

$$C_{ij} = \sum_{t=0}^{T} \left(R_{it} - \overline{R}_i\right) * \left(R_{jt} - \overline{R}_j\right) / T \tag{14.6}$$

We call Eq. (14.6) the historical model and will compare it with the factor model (Eq. 14.5). At the beginning of each period, previous 5 years' monthly returns are used to estimate betas of Eq. (14.4), factor model (Eq. 14.5), and historical model (Eq. 14.6).

Table 14.1 Global factor returns, February 2004–January 2014

Start date	Mkt	SMB	HML	WML
200402	12.77	7.42	8.68	3.75
200502	19.32	6.13	5.85	22.19
200602	17.00	−4.50	7.92	−0.70
200702	0.16	−8.19	−1.84	11.80
200802	−40.93	0.69	−5.13	19.53
200902	38.67	7.39	2.99	−40.31
201002	20.97	6.46	−0.76	13.07
201102	−3.64	−2.20	−8.42	2.19
201202	16.79	−4.19	9.19	10.56
201302	17.93	2.32	1.81	22.03
Average	9.90	1.13	2.03	6.41
Std	21.27	5.69	6.05	18.30
Corr				
Mkt	1.00	0.36	0.59	−0.48
SMB		1.00	0.18	−0.21
HML			1.00	−0.11
WML				1.00

There are numerous ways to estimate the expected return vector μ. Three groups of data have been shown to contain information of future returns. The first group of data consists of the raw returns, like past year's return, past 3 years' return, etc. The second group of data consists of risk-adjusted returns, like Treynor Index, Sharpe Ratio, and Jensen's alpha. This group of data measures the fund manager's stock selection skills by taking into account the portfolio's risk. The third group of data is the mutual fund's characteristics, like expense ratios, annual turnover rates, and top holdings concentrations. These three groups of data have been studied extensively in the literatures for US equity funds. We will review them in detail accordingly.

14.3.1 Raw Return

Can past performance of mutual funds be indicative of future performance? Hendricks et al. (1993) found strong evidence that last year's winners will continue to do well this year for US growth equity mutual funds using data from 1974 to 1988. Carhart (1997) studied the all equity mutual fund data from 1962 to 1993 and concluded that funds with the highest returns last year will have higher returns than average fund returns this year. Carlson (1970) using data from 1948 to 1967, and Brown and Goetzmann (1995) using data from 1970 to 1989, also found support of persistence of raw returns. However some researchers, like Brown et al. (1992), argue that the persistence is the result of survivor bias of the test database. In this study we show that the performance still persists even after control for survivor bias. Since the return from the manager's skill is small when compared to the return of risk factors, the underlying risk exposures and the persistence of risk factor returns are the main determinants of persistence of raw performance. If raw returns are used as the sole selection criteria for mutual fund, then high risk mutual funds are most likely to be recommended. If the risk factor returns reverse themselves, then last year's winners will perform poorly. Table 14.3 shows that the 2007s top winner decile portfolio underperformed the bottom loser decile portfolio by almost 4 % during the 2008 market crash. When the market reversed itself in 2009, the 2008s winner decile portfolio underperformed the loser decile portfolio by 9.04 %.

14.3.2 Risk-Adjusted Return

The risk-adjusted return is the standard performance measurement. The risk model has evolved from the single factor model to multifactor models like Eq. (14.4). Treynor (1965) is the first one to adjust raw returns to evaluate mutual fund performance. He created the Treynor Index, which is the raw return divided by the mutual fund's beta against market. Sharpe (1966) created the Sharpe ratio as the mutual fund performance measurement, which is the excess return divided by the standard deviation of the mutual fund's return. The Shape ratios based on the return and volatility from 1954 to 1963 are positively correlated with the Sharpe ratios calculated using the data from 1944 to 1953. Jensen (1969) used the regression alpha of fund returns to market returns to evaluate the performance of mutual funds. More recently, Carhart (1997) proposed the four factor model Eq. (14.4) to study the risk-adjusted returns using data from 1962 to 1993. He found the top decile portfolio based on previous years' returns did outperform the bottom decile. Most researchers, like Pastor and Stambaugh (2002a, b), Elton et al. (1996), Carhart (1997) found that the relative risk-adjusted performance persist from formation period to post formation period.

14.3.3 Mutual Fund Characteristics

The Index fund industry and academies have long argued that active fund managers can't beat the market on average because of the expenses. Kinnel (2010) reported that expense ratio is the most reliable predictor of mutual fund's future success. He sorted funds into quintiles by expense ratios and category, and found that least expensive fund group always out performs the most expensive fund group. Academies have studied the fund characteristics too. Carhart (1997) documented a negative effect for fees. Cremers and Petajisto (2009) found a negative effect of fund size on performance.

All the researchers mentioned above concluded that past relative performance can be used to forecast future relative performance. Based on the conclusion of previous studies, we will study the following information variables (Table 14.2).

Table 14.2 Definition of information variables

Alpha = regression alpha of Eq. (14.4) using previous 5 years' monthly returns. It is the long-term risk-adjusted return
ERatio = reported fund expense ratio for the previous year
LQRet = the previous quarterly return before the formation time
LY1Ret = the previous year's return before the formation time
LY3Ret = the previous 3 years' cumulative return before the formation time
TI1 = LY1Ret/β_{mp}, where β_{mp} is the market beta in Eq. (14.4) estimated by using previous 5 years' monthly returns. This is the modified 1-year Trenyor Index
TI3 = LY3Ret/β_{mp}, where β_{mp} is the market beta in Eq. (14.4) estimated by using previous 5 years' monthly returns. This is the modified 3-year Tranyor Index
SR1 = LY1Ret/σ_p, where σ_p is the monthly standard deviation estimated by using previous 5 years' monthly returns. This is the modified 1-year Sharpe ratio
SR3 = LY3Ret/σ_p, where σ_p is the monthly standard deviation estimated by using previous 5 years' monthly returns. This is the modified 3-year Sharpe ratio
Assets = the asset under management at the end of the previous year

14.4 Data and Universe

All the mutual fund data are from Morningstar Principia. Morningstar assigns each mutual fund to a category according to the fund's objective. Starting from January 2000, we download the monthly return, expense ratio, total net asset value, turnover ratio, and Morningstar ratings for all the mutual funds. At the end of January of each year, by that time the mutual fund's characteristic data is available, we will reconstruct our universe by considering only those mutual funds that have more than \$100 million assets under management. Elton et al. (1996) found that 1-year survival rate is 98% for fund with AUM more than 15 million. Our AUM cutoff makes our universe free of survivor bias. For the foreign large blend category, there are 117 funds with average expense ratio 1.28% at end of year 2003, and 264 funds with average expense ratio 0.90% at the year end of 2013. Since we need 5 year's return data to calculate alpha and betas in (Eq. 14.4), we further eliminate those mutual funds which don't have 5 years returns. The reported annual return is cumulative return from February of the report year to the January of the next year. The bottom decile return is calculated as the larger returns of the worst two decile portfolio returns. We do this because Carhart (1997) has shown that the difference in returns from the worst two decile portfolios is un-proportionally large when compared to the difference in returns from other adjacent decile portfolios. Table 14.3 reports the return differences.

Another way to look at the predictive power of variables is to examine the information coefficient (Table 14.4). The Table 14.4 presents the results for some of the variables listed in Table 14.2.

Table 14.3 Difference of top decile portfolio return to bottom decile portfolio return

Starting date	ERatio	Alpha	LQRet	LY1Ret	TI1	SR1	LY3Ret	TI3	SR3	Assets
200402	0.64	1.40	−0.40	2.71	2.47	3.08	3.68	3.40	3.41	−0.30
200502	−2.36	4.22	10.61	2.87	2.16	−6.32	3.78	0.36	−1.03	−0.27
200602	2.36	1.12	1.23	2.68	1.42	1.52	3.76	−1.22	−0.77	−2.00
200702	1.57	5.29	4.59	3.86	1.85	3.12	6.16	3.67	2.92	2.03
200802	1.25	2.05	1.34	−3.93	−1.30	−0.05	−4.63	−0.36	−3.28	3.15
200902	2.18	9.94	−11.76	−9.04	7.23	8.53	5.65	8.03	6.20	1.86
201002	0.34	2.50	4.75	4.41	3.04	2.98	1.47	1.42	1.61	0.06
201102	2.11	0.58	0.61	−2.82	−2.11	−1.94	0.40	0.37	0.26	−1.16
201202	1.96	0.23	1.32	−1.72	−0.76	1.00	−2.44	−2.47	−3.61	4.06
201302	−2.41	0.88	0.90	7.51	6.38	6.88	−0.05	0.23	−0.92	3.22
Average	0.76	2.82	1.32	0.65	2.04	1.88	1.78	1.34	0.48	1.07
Std	1.79	2.98	5.61	4.91	3.05	4.21	3.50	3.01	3.08	2.07
T	1.35	2.99	0.74	0.42	2.11	1.41	1.61	1.41	0.49	1.63

Table 14.4 Annual information ratios

Starting date	ERatio	Alpha	LY1Ret	LY3Ret	TI1	SR1	TI3	SR3
200402	0.23	0.21	0.29	0.48	0.38	0.33	0.49	0.49
200502	0.02	0.30	0.13	0.23	0.17	0.06	0.25	0.22
200602	0.27	0.00	0.02	0.24	−0.02	0.02	0.15	0.18
200702	0.20	0.15	0.03	0.35	−0.09	−0.03	0.26	0.28
200802	0.15	0.19	0.02	−0.08	0.10	0.11	0.06	0.06
200902	0.10	0.36	−0.02	0.13	0.14	0.22	0.14	0.15
201002	0.04	0.23	0.26	0.13	0.19	0.22	0.15	0.15
201102	0.20	0.26	0.08	0.30	0.12	0.13	0.28	0.28
201202	0.21	−0.12	0.01	−0.24	0.00	0.02	−0.26	−0.26
201302	−0.09	−0.03	0.49	0.10	0.45	0.44	0.07	0.07
Average	0.13	0.16	0.13	0.16	0.14	0.15	0.16	0.16
Std	0.11	0.15	0.15	0.20	0.16	0.14	0.18	0.18
T	3.93	3.34	2.70	2.61	2.84	3.37	2.72	2.79

Table 14.5 Distribution of minimum initial investment at 2013

Minimum initial investment	0	≤2500	≤5000	≥100,000
Count	30	39	4	52

The positive decile return differences and statistically significant positive information coefficients confirm that these variables are viable expected return models. The data period under consideration is a very special period. We experienced the great financial crises. The market tanked in 2008 and started to bounce back in 2009. Most variables failed in year 2009. In particular, the momentum variables (LQRet, LY1Ret, and LY3Ret) failed more than the risk-adjusted variables like TI3 and SR3.

14.4.1 Transaction Cost, Turnover, and Upper Bound

One of the decision variables of using mean-variance portfolio construction and management process (Eq. 14.1) is the turnover constraint from period to period. The optimal turnover depends on what expected return model to use and what transaction cost per trade the investor expects. The cost per trade is from 7 to 10 dollars for most online trading platforms. So the percentage cost of trading depends on the actual dollars amount traded. For example, 10 dollar fee for 2000 dollar trade translates into 50 basis points of cost, and 10 dollar fee for 4000 dollar trade translates into 25 basis points of cost. This makes higher upper bound a way of lowing transaction cost. For this reason we will run our simulation with two levels of transaction cost and two levels of upper bounds. The two levels of cost are 25 basis points and 0 basis points per trade. The two upper bounds are 10 and 20 %. There are other trade frictions too. One of them is the minimum initial investment. In order to implement the mean-variance portfolio weights, we remove those mutual funds with minimum initial investment more than 2500 dollars from our simulation universe. This deletion does not affect our simulation results. Another trade friction is the fee imposed by mutual funds for frequent traders. They are usually 2 and 1 % if the holding period of the fund is less than 3 and 6 months respectively. For this reason, we will simulate our portfolio construction process on an annual basis.

14.5 *Ex Post* Efficient Frontiers

We will run a series of simulations by varying the expected returns, turnover constraints, risk models, transaction cost, and upper bound. First we would like to settle what risk model to use. Table 14.6 compares the risk-return trade-off curves generated with factor risk model and historical risk model, using Alpha as expected returns with no transaction costs, no turnover constraints, and upper bounds of 10 %. There are no statistically significant differences for these two efficient sets. This is true for other expected return models. From now on we will report ex-post efficient frontiers using factor risk model only.

The next three exhibits validate our expected return models. Tables 14.7, 14.8, and Fig. 14.1 show the benchmark returns, and the efficient frontiers using different expected return models with no transaction cost, and no turnover constraints. The category average returns are reported by Morningstar. Index funds returns are calculated by the average return of the 10 least expensive funds, which turned out to be index funds. The long-term risk-adjusted return Alpha dominates the 1-year momentum variable LY1Ret, which dominates the fund expensive ratio variable ERatio. Alpha mean-variance efficient portfolios with comparable risks are 2 percentage points better than the benchmarks. LY1Ret and ERatio mean-variance efficient portfolios are as good as benchmarks.

Table 14.6 February, 2004–January, 2014, no cost, no turnover constraint

Expected return: Alpha						
Risk model	Factor			Historical		
λ	Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio
9	12.24	25.75	0.42	12.24	25.76	0.42
8	11.95	24.96	0.42	12.07	25.03	0.42
7	11.52	23.98	0.42	11.69	24.00	0.42
6	11.38	23.19	0.43	11.42	23.02	0.43
5	11.28	22.07	0.44	11.33	22.04	0.45
4	11.11	21.05	0.46	11.02	21.19	0.45
3	10.94	20.56	0.46	10.93	20.74	0.45
2	10.73	20.51	0.45	10.65	20.31	0.45
1	10.49	20.15	0.45	10.21	19.75	0.44

Table 14.7 Benchmark returns: February, 2004–January, 2014

Category average			Index funds		
Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio
9.28	23.37	0.33	8.81	21.24	0.34

Table 14.8 February, 2004–January, 2014, no cost, no turnover constraints

Expected return	LY1Ret			Alpha			ERatio		
λ	Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio
9	10.22	24.12	0.36	12.24	25.75	0.42	8.81	21.24	0.34
8	9.98	23.54	0.36	11.95	24.96	0.42	8.94	21.15	0.35
7	9.50	22.80	0.35	11.52	23.98	0.42	8.82	20.99	0.35
6	9.11	22.09	0.34	11.38	23.19	0.43	8.67	20.83	0.34
5	8.98	20.88	0.36	11.28	22.07	0.44	8.57	20.69	0.34
4	8.92	20.12	0.37	11.11	21.05	0.46	8.55	20.44	0.34
3	9.00	20.04	0.37	10.94	20.56	0.46	8.34	19.93	0.34
2	8.79	20.02	0.36	10.73	20.51	0.45	8.41	19.46	0.35
1	8.90	19.13	0.39	10.49	20.15	0.45	8.61	19.08	0.37

The *ex post* efficient frontier is not as smooth and as concave as ex ante efficient frontier. The main cause is the imperfect estimation of expected returns. Nevertheless, the portfolios with less risk-tolerant parameter, i.e., lower λ_E in ex-ante, do realize less total risk in *ex post*. The mean-variance process is very effective on controlling portfolio risk. In particular, the minimum variance portfolio trades off only 2 % return with 6 % less risk when compared to the return-maximizing portfolio for Alpha model. The conservative portfolios have higher Sharpe ratios than the more risk taking portfolios. The performance of expensive ratio is not as good as expected since it has an average positive IC of 0.13 and highest *t* statistic of 3.93. On the other hand, the least expensive funds are dominated by index funds so one should not expect them to outperform the index fund benchmark. The positive

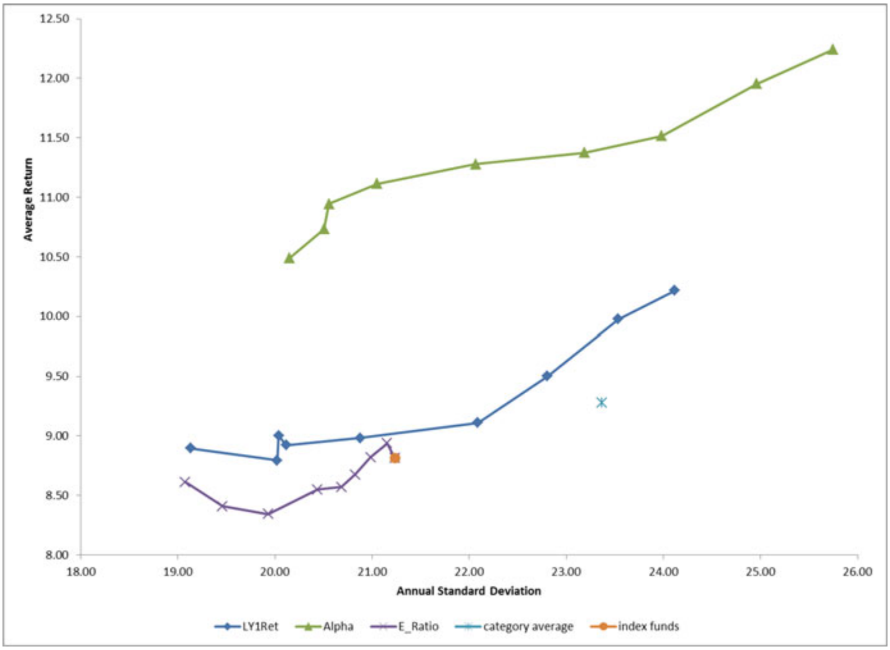


Fig. 14.1 February, 2004–January, 2014, no cost, no turnover constraints

decile portfolio spread and positive IC suggest that we should either combine it with other variables or use it as a constraint to avoid the most expensive funds.

In the presence of transaction cost, we expect the efficient portfolios to underperform no cost-efficient ones by

$$2 \times \text{Cost} \times \text{Turnover} \tag{14.7}$$

Table 14.9 shows the cost effects on momentum LY1Ret model and Alpha model. The performance deduction on LY1Ret model confirms to Eq. (14.7). The resulting portfolios underperform the benchmarks. However, there are no performance deductions for Alpha model. It is possible because the optimizer takes into account the transaction cost, and the portfolios are different. The portfolios with transaction cost turned out to be better for the Alpha model even after the cost.

On one hand, imposing turnover constraints puts an upper bound (Eq. 14.7) on trading cost. On the other hand, it may prevent the optimizer from fully utilizing the predictive information. Table 14.10, Figs. 14.2 and 14.3 show the overall effect of turnover and cost. Adding turnover constraints, the LY1Ret model performs almost 1 % better while Alpha model performs 1 % worse respectively.

All the above simulations are done with 10 % as the upper bound for each position to enforce some diversification. Table 14.11 and Figs. 14.4 and 14.5 show the efficient frontiers by loosening the upper bound from 10 to 20 %. For Alpha

Table 14.9 February, 2004–January, 2014, cost = 0.25%, no turnover constraints

Expected return	LY1Ret			Alpha		
λ	Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio
9	9.69	24.09	0.34	12.18	25.82	0.41
8	9.41	23.21	0.34	11.96	24.56	0.43
7	8.92	22.97	0.32	12.05	23.57	0.45
6	8.73	21.82	0.33	11.30	22.58	0.43
5	9.01	20.83	0.36	11.42	21.90	0.45
4	8.83	20.21	0.36	11.06	20.97	0.46
3	8.90	20.01	0.37	10.81	20.46	0.46
2	8.57	20.05	0.35	10.61	20.49	0.44
1	8.75	19.24	0.38	10.31	20.26	0.43

Table 14.10 February, 2004–January, 2014, cost = 0.25, turnover ≤ 0.25

Expected return	LY1Ret			Alpha		
Λ	Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio
9	11.17	25.36	0.38	11.50	25.04	0.40
8	10.34	22.85	0.39	10.40	22.55	0.39
7	10.08	21.66	0.40	10.30	21.64	0.41
6	10.24	21.51	0.41	10.28	21.50	0.41
5	10.13	20.73	0.42	9.79	20.79	0.40
4	9.64	20.38	0.40	10.70	20.35	0.45
3	9.24	19.56	0.40	10.40	20.28	0.44
2	9.02	19.22	0.39	10.13	20.16	0.43
1	8.73	19.03	0.38	9.68	19.72	0.41

model, the whole efficient frontier shifts to upper-left more than 1 %. The return is more and risk is less. For LY1Ret model, the effect is mixed. The riskier portfolios underperform and the conservative portfolio outperform than the tighter constrained portfolios.

14.5.1 Conclusion

This paper discusses the nontrivial aspects of applying the mean-variance optimization technique to manager portfolios of foreign large blend mutual funds. There are numerous viable expected return models from past performance and fund characteristics. With the right turnover constraints, upper bound, and appropriate risk-return trade-off parameter, efficient portfolios with comparable risk as the benchmarks can outperform the benchmarks more than 2 %.

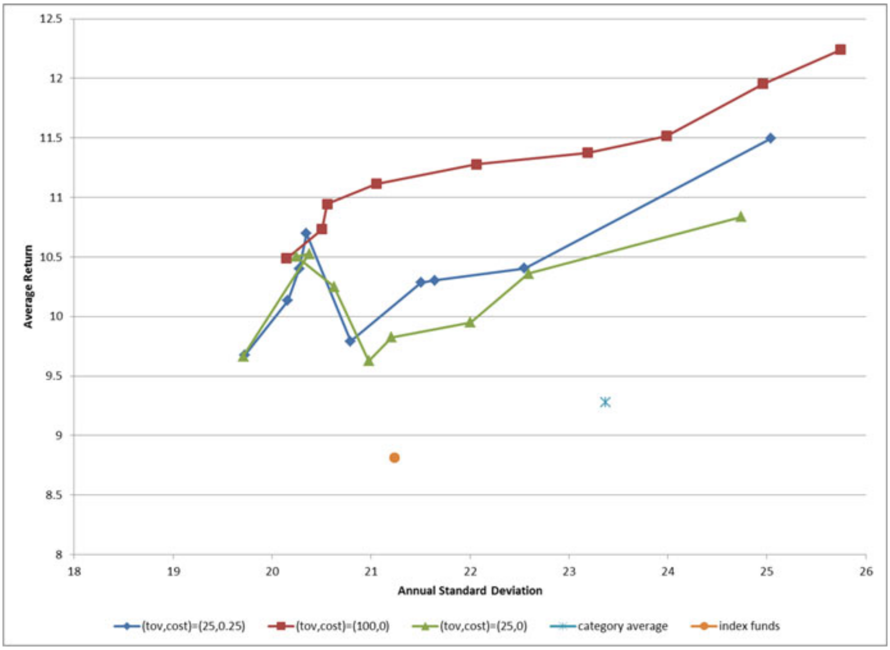


Fig. 14.2 Turnover and cost effect of Alpha model

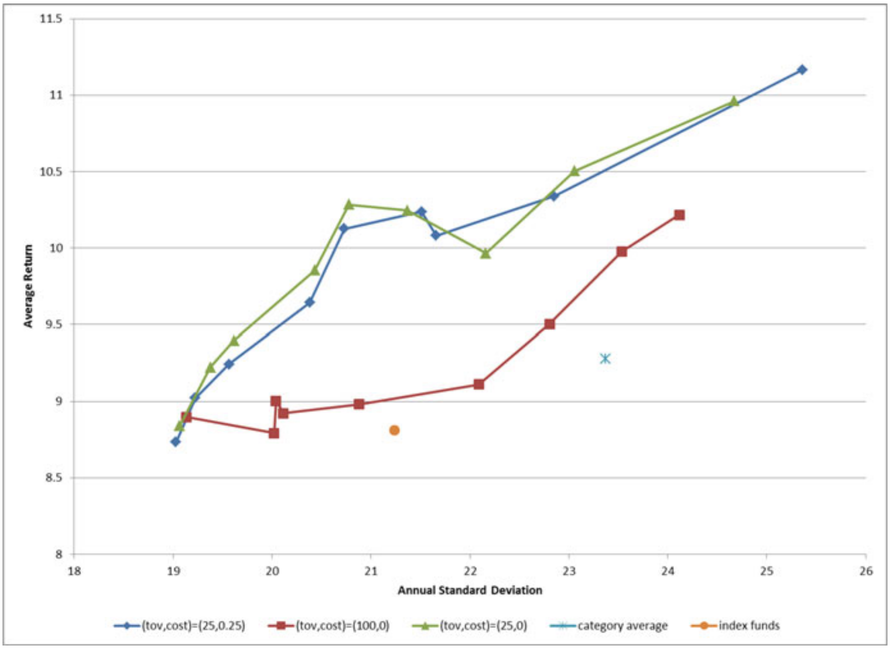


Fig. 14.3 Turnover and cost effect of LY1 model

Table 14.11 February, 2004–January, 2014, cost = 0.0, turnover≤25 %, upper bound≤0.20

Expected return	LY1Ret			Alpha		
λ	Mean	Std	Sharpe ratio	Mean	Std	Sharpe ratio
9	10.25	25.01	0.35	11.32	23.74	0.41
8	9.36	22.65	0.35	10.73	21.27	0.43
7	9.26	21.44	0.36	11.33	20.23	0.49
6	9.92	20.50	0.41	10.80	18.96	0.49
5	10.23	19.53	0.45	10.02	18.66	0.46
4	9.72	19.31	0.43	10.18	18.78	0.46
3	9.38	18.41	0.43	9.74	18.56	0.44
2	9.09	17.57	0.43	9.88	18.27	0.46
1	9.21	17.36	0.44	10.02	18.35	0.46

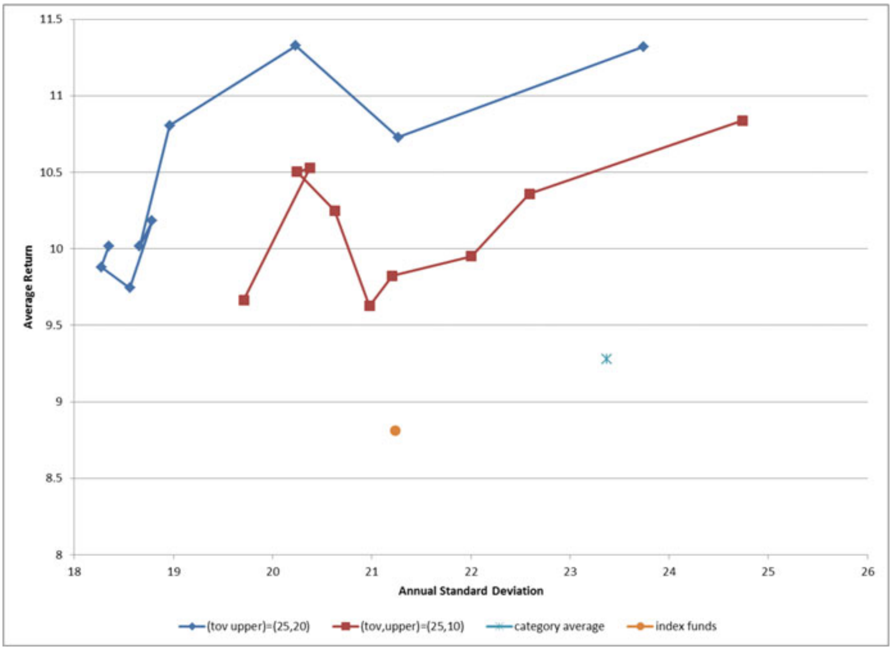


Fig. 14.4 Alpha model with cost = 0.0, turnover≤25 %, upper bound≤0.20

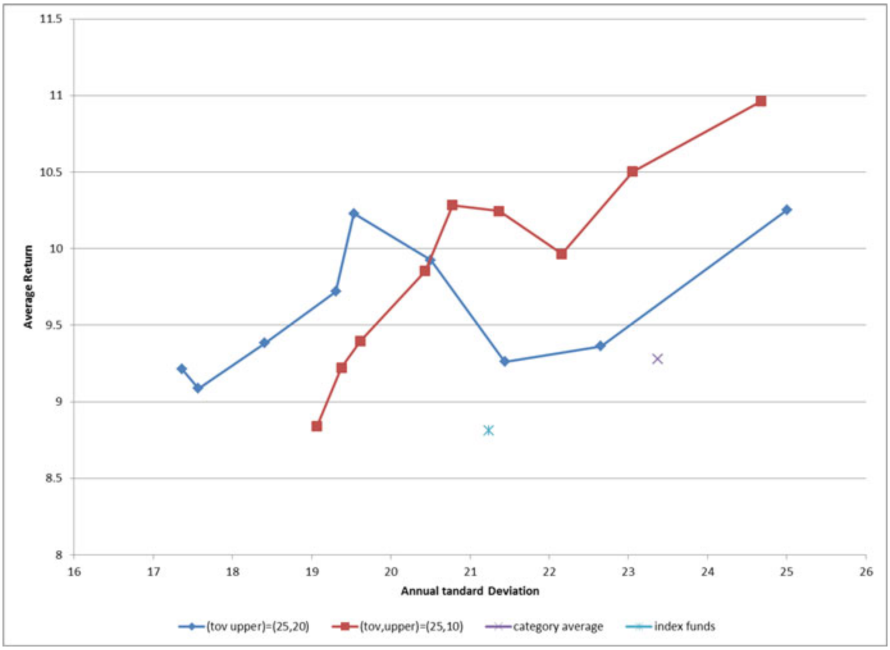


Fig. 14.5 LY1Ret model with cost = 0.0, Turnover ≤ 25 %, upper bound ≤ 0.20

Acknowledgements The authors would like to thank Dianne Toombs, Chunrong Pan for preparing the data used in this study.

References

Brown, S., & Goetzmann, W. (1995). Performance persistence. *Journal of Finance, 50*(2), 679–698.

Brown, S., Goetzmann, W., Ibbotson, R., & Ross, S. (1992). Survivorship bias in performance studies. *Review of Financial Studies, 5*(4), 553–580.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(l), 57–82.

Carlson, R. S. (1970). Aggregate performance of mutual funds: 1948–1967. *Journal of Financial and Quantitative Analysis, 5*(1), 1–32.

Cremers, M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies, 22*, 3329–3365.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). The persistence of risk-adjusted mutual fund performance. *Journal of Business, 69*, 133–157.

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot hands in mutual funds: Short-run persistence of relative performance. *Journal of Finance, 48*, 93–130.

Jensen, M. (1969). Risk, the pricing of capital assets and the evaluation of investment portfolios. *Journal of Business, 42*, 167–247.

Fama, E. F., & French, K. R. (1992). Cross-sectional variation in expected stock returns. *Journal of Finance, 47*(2), 427–465.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*(1), 55–84.

French, K. R. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investment.* Cowles Foundation Monograph No. 16. New York: John Wiley & Sons.

Pastor, L., & Stambaugh, R. F. (2002a). Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics, 61*, 315–349.

Pastor, L., & Stambaugh, R. F. (2002b). Investing in equity mutual funds. *Journal of Financial Economics, 63*, 351–380.

Kinnel, R. (2010). How expense ratios and star rating predict success. August 9, Morningstar.com

Sharpe, W. (1966). Mutual fund performance. *Journal of Business, 39*, 119–138.

Treynor, J. (1965). How to rate management of investment funds. *Harvard Business Review*, January–February, pp. 63–75.

Chapter 15
Fundamental Versus Traditional Indexation for International Mutual Funds: Evaluating DFA, WisdomTree, and RAFI PowerShares

Heehyun Lim and Edward Tower

The fundamental index fund is a hybrid of active and passive management. Fundamental indexation is passive in that it uses rules for portfolio selection. It is active in the sense that its portfolio weights deviate from market cap weights, and therefore fundamental indexation takes an active position relative to the cap-weighted market.

Do fundamental index funds beat traditional ones? The major companies that offer the new fundamental index international mutual funds are Dimensional Fund Advisors (DFA), Research Affiliates, and WisdomTree. A major provider of traditional international index funds is DFA. We compare various fundamental index fund portfolios from these companies with individualized benchmark portfolios composed of DFA traditional funds.

Jeremy Siegel said in an interview with the *New York Times* in 2006, traditional index funds overweight overvalued stocks while they underweight undervalued stocks, causing investors to buy fashionable assets at high prices (Anderson, 2006). Robert Arnott, Chairman of Research Affiliates, succinctly describes fundamental indexation (PowerShares, 2012 p. 3): "Fundamental Index strategies use Fundamentals, various measures of firm size, including dividends, earnings, cash flow,

Edward Tower is a Professor of Economics at Duke University and a long-term visiting professor at Chulalongkorn University in Bangkok Thailand.
Heehyun Lim is a Phd Candidate and Teaching Fellow at the University of Maryland, College Park.

H. Lim (✉)
University of Maryland, College Park, MD, USA
e-mail: heehyun.rosa.lim@gmail.com

E. Tower
Economics Department, Duke University, Durham, NC, USA
e-mail: tower@econ.duke.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_15

sales, book-equity values, and so on." For more complete descriptions of active and fundamental index strategies see Arnott and Fabozzi (1992) and Arnott et al. (2008).

The opponents of fundamental index funds, however, claim that the excess return generated from fundamental indexation is minor and that the additional costs, including turnover costs, and tax inefficiencies, cancel out the advantage of this alternative index fund, if there is any.

Tower and Yang (2013) compare DFA, Research Affiliates, and WisdomTree US fundamental index funds with portfolios constructed out of Vanguard index funds. Here, by examining international funds we complete the inquiry that Tower & Yang initiated.

We use a modified "style analysis" of Sharpe (1992) to compare the performance of traditional index funds and fundamental index funds of international stocks. Vanguard has a limited range of international index funds, so instead of using Vanguard traditional index funds to benchmark, here we employ DFA traditional index funds to do the job.

Tower and Yang (2013) compare Vanguard, a traditional index fund manager, with leading fundamental index fund families, DFA-Core and Vector funds, Power shares-Research Affiliates (RAFI for short), and WisdomTree. They use two Fama-French models as well as Sharpe's style analysis (Sharpe 1992). They find that the RAFI funds and WisdomTree funds outreturned their Vanguard clones, whereas DFA core slightly underreturned. Instead of supporting either fundamental or traditional funds, their study produces the conflicted result: "two cheers for enhanced [or more properly fundamental] indexation and one for traditional." Their results for the Fama-French models and the Sharpe style analysis models were quite similar. So here we simplify matters by using only the Sharpe (1992) style analysis.

To clarify terminology, one set of readers (identified in the acknowledgements) wrote, "Enhanced index funds are the standard term used in the finance industry to denote funds which manage off a given benchmark index, usually with an implicit or explicit tracking error constraint. RAFI and WT index funds do not fall into this category. We would suggest using the term 'smart beta,' 'alternative index' or 'strategic index,' instead." There are advocates for all three. We choose the term "fundamental index funds."

15.1 Style Analysis

Sharpe (1992) introduces a useful way of comparing fund performance. He constructs a clone portfolio of indexes that minimizes the variance of the difference between the return on the fund and the synthetic portfolio. The clone portfolio is a combination of indexes that reflects the manager's asset mix. For example, if a manager maintains a portfolio consisting of 25 % small value stocks and 75 % large growth stocks, running a regression of the return of the mutual fund on the returns of a set of indexes, not suppressing the constant term, and constraining the sum of the portfolio weights, represented by the regression coefficients, to add up to 1, should yield a coefficient of 0.25 for the small value index and 0.75 for the large growth index. Lucas and Riepe (1966) also provide a transparent discussion of returns-based style analysis.

We use a modified version of style analysis as Tower (2009) more comprehensively explains, replacing indices with index funds. Using 11 DFA international traditional index funds, we create the clone portfolio as that which best mimics the pattern of returns of each individual international fundamental fund or portfolio of fundamental funds, henceforth, simply fundamental fund.

15.2 How Do We Create the Clone Portfolio?

We instruct Microsoft Excel Solver to find the weighted sum of each traditional fund's return to produce the fundamental fund's return. By allowing the weights, which must sum up to 1, to vary, we can find the particular set of weights that minimizes the standard deviation of return differentials between the fund and its clone portfolio. In short, solver finds the set of coefficients that minimizes the variance of the return differentials between the fundamental fund and the traditional fund portfolio. Thus we regress the return of the fundamental fund on the returns of the traditional funds, while constraining the portfolio weights to sum to 1. We assume that the managers of the fundamental index fund portfolio and the clone portfolio rebalance once each month in order to maintain the portfolio weights.

15.3 Why Use a Clone Portfolio and Style Analysis?

We compare each fundamental fund portfolio with the collection of DFA traditional funds which tracks it best. The geometric average excess return of the fundamental fund, continuously compounded, is α. It represents how much value has been added by fundamental indexation.

Our readers noted "Since the clone portfolio uses traditional indexes that vary across regions (like Asia, Europe, Japan, developed and emerging markets, etc.) and across sizes (small cap, large cap) and styles (value and growth), measuring an index against its clone will remove all outperformance or underperformance due to region, size and style. The remaining alpha, therefore is the "skill" that must be due to something other than picking the right region, size or style."

"It is important to note that the clone portfolios are created *ex post*. So even if a manager has tremendous foresight and can pick the region/size/style combination that will perform the best over the sample period, he is evaluated relative to a clone portfolio that has the benefit of picking that region/size/style after the fact. ... [Thus we] "are looking for skill that goes beyond this, such as picking stocks within these buckets or dynamically shifting allocation among these buckets in a way that generates alpha."

15.4 Why Use Continuous Compounding and Geometric Average Return?

In order to measure the average rate of return over multiple time periods, we employ continuously compounded geometric average return. A portfolio that returns 50 % 1 year and −50 % the next year does not have an average return of zero % per year. But a portfolio that returns a continuously compounded 50 % in 1 year and −50 % in the next year does have a geometric average return of 0 % over the 2 years. Thus, the geometric average return over a span is the average of the geometric average returns over the periods that comprise the span.

15.5 Why Use Equally Weighted Portfolios and Risk-Averse Portfolios?

To compare the performance of traditional indexation and fundamental indexation, we consider two types of portfolio composed of fundamental index funds from each fund family. This approach is more relevant than looking at individual funds, because typical investors hold a variety of funds. Thus, in addition to looking at each fund separately, we create an equally weighted portfolio and a risk-averse portfolio to compare the returns with a traditional clone portfolio. While the equally weighted portfolio is not necessarily optimal, it is by far the simplest way to look at the overall performance of the assets in a fund family.

The risk-averse portfolio, however, is perhaps more useful, especially when we account for the investment atmosphere after the financial crisis, with the obvious need to make safe decisions, and we recognize that some funds, such as WisdomTree's India Earnings Fund, are highly specialized and would constitute a small proportion of a diversified portfolio. Therefore, in order to reflect the preference for less risk, we create a risk-averse portfolio of each fund family. The weights of each risk-averse fundamental index fund portfolio were determined to minimize the standard deviation of the return of the portfolio, again using Microsoft Excel Solver.

15.6 Why Do Some of Our Portfolios Allow Short Selling?

The Sharpe style analysis constrains the portfolio weights to be nonnegative. In other words, it does not allow holding short positions in any fund to create a clone portfolio. However, when an investor who initially held traditional funds buys a fundamental fund, in some cases, he needs to increase his holdings of some traditional funds to imitate his previous style.

For example, suppose the fundamental fund is more focused on growth companies than any of the DFA traditional funds. Our investor has an initial portfolio of DFA traditional funds. She invests a dollar in the fundamental fund. She maintains her style by selling two dollars of her DFA traditional blend fund and buying one dollar worth of the DFA traditional value fund. Then the clone portfolio is 200 % DFA traditional blend and −100 % DFA traditional value. Regressing the fundamental fund return on the traditional fund returns, while constraining the coefficients to sum to one and not repressing the constant term, would yield a coefficient of 2 for the DFA traditional blend fund and −1 for the DFA traditional value fund.

Therefore, a negative coefficient for a particular fund in the clone portfolio signals that the funds with positive coefficients in the clone portfolio are leveraged in order to achieve the same style as before, and the investor buys more of the traditional fund with the negative coefficient to maintain portfolio balance when she buys the fundamental fund. The coefficients show the net sales of each traditional fund necessary to maintain style when a dollar's worth of the traditional fund is purchased. While traditional index funds cannot be sold short, ETFs can be sold short, and some ETFs mimic traditional index funds.

15.7 Data

The ideal pick of funds is mutually exclusive but exhaustive (Sharpe, 1992). The data were selected to conform as much as possible to these criteria by eliminating the funds with redundant components. For example we only use one of the DFA International Value funds. The data on monthly returns were collected from the Center for Research in Security Prices (CRSP), and some missing data were filled from Yahoo Finance and Morningstar. In this study, the portfolios of fundamental DFA funds have monthly returns spanning from between May 2005 and September 2008 to June 2014, those for RAFI span from between July 2007 and October 2007 to June 2014, and those for WisdomTree span from between July 2006 and December 2009 to June 2014. We examined all the fundamental international stock mutual funds from the three fund families that existed during our time frame except for sector funds.

15.8 The Exhibits

To construct the clone portfolios, we use 11 DFA traditional index funds. They are listed alphabetically by ticker next to the bottoms of Exhibits 15.1, 15.2, 15.3, 15.4, 15.5, and 15.6. In all the Exhibits,

Exhibit 15.1 DFA fundamental index funds. No shorts

Ticker and name All figures in %	DFCCX CSTG&E Intl Social Core Equity	DFCEX Emerging Markets Core Equity I	DFESX Emerging Markets Social Core Equity	DFIEX Intl Core Equity I	DFSPX Intl Sustainability Core I	DFVQX Intl Vector Equity Inst I	**DFA Equal Weights**	**DFA Risk Averse. 100 % in DFSPX**	**Avg for individual funds**	**Median for individual funds**
Inception date	Aug-07	Apr-05	Aug-06	Sep-05	Mar-08	Aug-08	Aug-08	Mar-08	Dec-06	Feb-07
α. %/year	**0.14**	**−0.24**	**−0.41**	**−0.27**	**−0.53**	**0.15**	**−0.06**	**−0.53**	**−0.19**	**−0.26**
Excess St.D of fund. %/month	0.006	−0.043	0.040	0.013	0.020	−0.019	−0.009	0.02	0.003	0.010
St.D of predict. Error. %/month	0.24	0.29	0.31	0.18	0.22	0.30	0.16	0.22	0.24	0.27
Correlation	99.93	99.92	99.92	99.95	99.94	99.91	99.97	99.94	99.94	99.93
Significance one tail	33	23	15	10	5	27	40	5	22	19
DEMSX Emg Mkts Small Cap	0	22	18	0	1	0	8	1	7	1
DFALX Large Cap Intl	38	1	0	38	53	2	20	53	22	20
DFCSX Continental Small Co	4	0	0	1	4	0	1	4	2	0
DFEMX Emerging Markets	0	43	44	1	0	2	17	0	15	2
DFEVX Emg Mkts Value	0	32	36	0	0	0	9	0	11	0

DFISX Intl Small Co	0	0	0	6	8	10	1	8	4	3
DFIVX International Value	34	1	0	29	20	45	24	20	22	25
DFJSX Japanese Small Co	4	0	0	2	2	3	2	2	2	2
DFRSX Asia Pacific Small Co	3	0	1	3	0	4	3	0	2	2
DFUKX U.K. Small Co	1	1	0	3	2	3	3	2	2	2
DISVX Intl Small Cap Value	15	0	0	18	8	29	12	8	12	11
Sum of shorts	0	0	0	0	0	0	0	0	0	0
α 1st half. %/year	0.77	−0.51	−0.26	−0.51	−0.62	0.59	0.28	−0.62	−0.09	−0.39
α 2nd half. %/year	−0.70	0.01	−0.39	0.01	−0.58	−0.08	−0.24	−0.58	−0.29	−0.24
α average. %/year	**0.04**	**−0.25**	**−0.33**	**−0.25**	**−0.60**	**0.26**	**0.02**	**−0.60**	**−0.19**	**−0.31**

Exhibit 15.2 DFA fundamental index funds. Shorts permitted

Ticker and name All figures in %	CSTG&E Intl Social Core Equity	Emerging Markets Core Equity I	Emerging Markets Social Core Equity	Intl Core Equity I	Intl Sustainability Core I	Intl Vector Equity Inst I	**DFA Equal Weights**	**DFA Risk Averse. 100 % in DFSPX**	**Avg for individual funds**	**Median for individual funds**
Inception date	Aug-07	Apr-05	Aug-06	Sep-05	Mar-08	Aug-08	Aug-08	Mar-08	Dec-06	Feb-07
Alpha. %/year	**0.14**	**−0.27**	**−0.49**	**−0.27**	**−0.65**	**0.22**	**−0.06**	**−0.65**	**−0.22**	**−0.27**
Excess St.D of fund. %/month	0.006	−0.028	0.014	0.013	0.022	−0.019	−0.009	0.02	0.001	0.010
St.D of predict. Error. %/month	0.23	0.28	0.22	0.18	0.22	0.30	0.16	0.22	0.24	0.23
Correlation	99.94	99.93	99.92	99.95	99.95	99.91	99.97	99.95	99.93	99.93
Significance one tail	32	20	10	10	2	30	40	2	17	15
DEMSX Emg Mkts Small Cap	−4	22	20	0	6	−2	8	6	7	3
DFALX Large Cap Intl	39	−1	−6	38	52	1	20	52	21	19
DFCSX Continental Small Co	8	0	−2	1	2	0	1	2	1	0
DFEMX Emerging Markets	2	47	49	1	0	3	17	0	17	3
DFEVX Emg Mkts Value	2	29	30	0	−5	1	9	−5	9	2

DFISX Intl Small Co	−10	−9	0	6	13	11	1	13	2	3
DFIVX International Value	32	4	5	29	22	45	24	22	23	26
DFJSX Japanese Small Co	6	2	−1	2	1	3	2	1	2	2
DFRSX Asia Pacific Small Co	5	3	2	3	0	5	3	0	3	3
DFUKX U.K. Small Co	3	4	0	3	1	3	3	1	2	3
DISVX Intl Small Cap Value	16	−2	4	18	9	28	12	9	12	12
Sum of shorts	−14	−12	−10	0	−6	−2	0	−6	−7	−8
α 1st half. %/year	−1.65	−0.64	−0.34	−0.53	−0.73	1.39	0.28	−0.73	−0.42	−0.59
α 2nd half. %/year	−0.87	−0.07	−0.35	−0.01	−0.74	0.03	−0.24	−0.74	−0.34	−0.21
α average. %/year	**−1.26**	**−0.36**	**−0.35**	**−0.27**	**−0.74**	**0.71**	**0.02**	**−0.74**	**−0.38**	**−0.35**

Exhibit 15.3 RAFI. Fundamental index funds. No Shorts

Ticker and name All figures in %	PAF FTSE RAFI Asia Pacific ex-Japan	PDN FTSE RAFI Developed Market ex-US Small-Mid	PXF FTSE RAFI Developed Markets ex-US	PXH FTSE RAFI Emerging Market	**RAFI Equal Weights**	**RAFI Risk Averse. 94.3 % PDN&5.7 % PXF**	**Avg for individual funds**	**Median for individual funds**
Inception date	Jun-07	Sep-07	Jun-07	Sep-07	Sep-07	Sep-07	Jul-07	Jul-07
α. %/year	**2.87**	**0.30**	**0.17**	**-1.93**	**0.66**	**0.29**	**0.35**	**0.24**
Excess St.D of fund. %/month	0.09	−0.01	−0.02	−0.02	−0.12	−0.05	0.01	−0.02
St.D of predict. Error. %/month	1.73	0.94	2.67	1.21	0.99	0.95	1.64	1.47
Correlation	97.24	98.91	92.38	98.72	98.98	98.90	96.81	97.98
Significance one tail	11	42	48	12	31	40	28	27
DEMSX Emg Mkts Small Cap	0	18	6	0	0	17	6	3
DFALX Large Cap Intl	16	0	22	3	10	0	10	9
DFCSX Continental Small Co	0	4	10	0	0	0	4	2
DFEMX Emerging Markets	62	0	3	75	51	0	35	33
DFEVX Emg Mkts Value	0	0	0	22	0	0	6	0
DFISX Intl Small Co	0	7	0	0	0	11	2	0
DFIVX International Value	0	7	50	0	21	11	14	4
DFJSX Japanese Small Co	0	16	1	0	1	14	4	0
DFRSX Asia Pacific Small Co	22	0	0	0	2	0	6	0
DFUKX U.K. Small Co	0	12	9	0	3	11	5	5
DISVX Intl Small Cap Value	0	35	0	0	12	36	9	0
Sum of shorts	0	0	0	0	0	0	0	0
α 1st half. %/year	3.79	2.27	0.35	0	1.85	2.16	1.60	1.31
α 2nd half. %/year	3.76	0.14	2.98	−3.18	0.87	0.22	0.93	1.56
α average. %/year	**3.78**	**1.21**	**1.665**	**−1.59**	**1.36**	**1.19**	**1.26**	**1.44**

Exhibit 15.4 RAFE. Fundamental index funds. Shorts permitted

Ticker and name	PAF FTSE RAFI Asia Pacific ex-Japan	PDN FTSE RAFI Developed Market ex-US Small-Mid	PXF FTSE RAFI Developed Markets ex-US	PXH FTSE RAFI Emerging Market	**RAFI Equal Weights**	**RAFI Risk Averse. 94.3 % PDN&5.7 % PXF**	**Avg for individual funds**	**Median for individual funds**
Inception date	Jun-07	Sep-07	Jun-07	Sep-07	Sep-07	Sep-07	Jul-07	Jul-07
α. %/year	**2.77**	**0.21**	**-0.13**	**-0.48**	**0.47**	**0.22**	**0.59**	**0.04**
Excess St.D of fund. %/month	0.07	−0.05	0.08	−0.02	−0.11	−0.07	0.02	0.02
St.D of predict. Error. %/month	1.47	0.93	2.61	1.09	0.98	0.94	1.52	1.28
Correlation	98.02	98.94	92.63	98.97	98.99	98.92	97.14	98.48
Significance one tail	8	43	49	37	35	43	34	40
DEMSX Emerging Markets Small Cap I	−13	25	16	−17	0	24	3	2
DFALX Large Cap International I	−10	−23	46	30	11	−18	11	10
DFCSX Continental Small Company I	17	−3	19	−22	7	0	3	7
DFEMX Emerging Markets I	150	8	3	66	55	8	57	37
DFEVX Emerging Markets Value I	−74	−17	7	54	−3	−15	−8	−5
DFISX International Small Company I	−83	35	−37	23	−23	29	−16	−7
DFIVX International Value I	26	25	35	−3	19	26	21	25
DFJSX Japanese Small Company I	8	13	1	−10	6	13	3	4
DFRSX DFA Asia Pacific Small Compay I	53	0	−19	−16	6	−1	4	−8
DFUKX United Kingdom Small Co I	5	5	14	−9	6	6	4	5
DISVX International Small Cap Value I	23	30	15	5	17	29	18	19
Sum of shorts	−180	−42	−56	−78	−27	−35	−89	−67
α 1st half. %/year	3.28	2.87	1.53	2.02	1.88	2.82	2.82	2.45
α 2nd half. %/year	3.56	0.87	5.99	0	2.07	1.10	1.10	2.22
α average. %/year	**3.42**	**1.87**	**3.76**	**1.01**	**1.98**	**1.96**	**2.65**	**2.65**

Exhibit 15.5 WISDOMTREE. Fundamental index funds. No shorts

Ticker and name All figures in %	AUSE Australia Dividend	AXJL Asia Pacific ex-Japan	DEM Emg Mkts Equity Income	DFE Europe SmallCap Dividend	DFJ Japan SmallCap Dividend	DGS Emg Mkts SmallCap Divdnd	DIM Intl MidCap Dividend	DLS Intl SmallCap Dividend	DNL Global ex-US Growth	DOL Intl LargeCap Dividend
Inception date	Jun-06	Jun-06	Jul-07	Jun-06	Jun-06	Oct-07	Jun-06	Jun-06	Jun-06	Jun-06
α. %/year	**0.71**	**0.80**	**1.45**	**0.41**	**−0.74**	**0.19**	**−0.22**	**−0.55**	**−1.29**	**−0.20**
Excess St. D of fund %/month	0.30	−0.04	−0.36	0.27	−0.22	−0.30	−0.01	−0.02	−0.26	−0.06
St. D. of predict error %/mo	2.56	1.67	1.79	1.49	1.12	1.68	0.90	1.06	2.87	1.07
Correlation	93.9	96.6	96.7	97.9	96.7	97.5	98.9	98.4	83.8	98.3
Significance one tail	41	35	27	41	30	47	42	34	36	44
DEMSX Emg Mkts Small Cap	0	0	0	0	0	60	7	6	0	0
DFALX Large Cap Intl	17	22	0	0	0	0	35	0	73	98
DFCSX Continental Small Co	0	0	0	64	0	0	24	30	0	2
DFEMX Emerging Markets	12	56	80	0	0	25	0	0	1	0
DFEVX Emg Mkts Value	0	0	0	0	0	0	0	0	0	0
DFISX Intl Small Co	0	0	0	0	0	0	0	17	0	0
DFIVX International Value	11	0	0	0	0	0	10	0	0	0
DFJSX Japanese Small Co	3	8	9	0	96	4	11	21	23	0
DFRSX Asia Pacific Small Co	51	15	0	0	0	0	0	7	0	0
DFUKX U.K. Small Co	0	0	11	36	4	12	13	18	3	1
DISVX Intl Small Cap Value	6	0	0	0	0	0	0	0	0	0
Sum of shorts	0	0	0	0	0	0	0	0	0	0
α 1st half. %/year	3.66	0.72	6.49	−2.64	−0.45	3.18	−1.14	−0.09	2.22	−0.48
α 2nd half. %/year	−0.61	0.74	−1.11	5.94	−0.62	−1.49	−0.82	0.68	−1.39	0.04
α average. %/year	**1.53**	**0.73**	**2.69**	**1.65**	**−0.54**	**0.85**	**−0.98**	**0.30**	**0.42**	**−0.22**

DOO Intl Divend ex-Financials	DTH DEFA Equity Income	DWM DEFA	DXJ Japan Dividend	EPI India Earnings	GULF Middle East Dividend	HEDJ European Hedged Equity	**WT Equal Weights**	**WT Risk Averse. 75 % DNL & 25 % DFJ**	**Avg for individual funds**	**Median for individual funds**
Jun-06	Jun-06	Jun-06	Jun-06	Feb-08	Jul-08	Nov-09	Feb-08	Jun-06	Jan-07	Jun-06
−0.27	**−0.62**	**−0.14**	**−2.50**	**−5.13**	**−6.33**	**−0.02**	**−0.94**	**−1.08**	**−0.85**	**−0.22**
0.02	0.14	1.39	0.32	1.68	0.82	−0.38	−0.16	−0.44	0.22	0.01
1.32	1.43	3.46	2.53	4.91	6.33	5.01	0.91	2.18	2.51	1.74
97.7	97.7	80.2	84.9	87.8	31.7	6.4	98.9	88.9	81.9	96.7
43	35	49	22	23	24	50	23	34	35	35
1	0	0	0	48	0	0	10	0	7	0
55	69	0	32	0	0	6	28	54	22	3
8	9	0	0	0	0	0	11	0	8	0
2	0	6	0	0	0	15	23	1	13	1
0	0	0	0	52	0	0	0	0	3	0
0	0	0	0	0	0	0	0	0	2	0
30	19	0	0	0	0	0	0	0	4	0
0	0	72	67	0	60	79	18	41	27	10
0	0	0	0	0	0	0	0	0	5	0
5	4	21	1	0	40	0	10	4	9	4
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
−0.84	−1.89	−0.19	−3.60	−9.47	−18.41	−10.67	−1.74	1.63	−1.98	−0.48
−0.33	1.06	0.25	−1.40	−3.65	6.16	11.02	0.66	−1.09	0.85	−0.33
−0.59	**−0.42**	**0.03**	**−2.50**	**−6.56**	**−6.13**	**0.18**	**−0.54**	**0.27**	**−0.56**	**−0.41**

Exhibit 15.6 WISDOMTREE. Fundamental index funds. Shorts permitted

Ticker and name All figures in %	AUSE Australia Dividend	AXJL Asia Pacific ex-Japan	DEM Emg Mkts Equity Income	DFE Europe SmallCap Dividend	DFJ Japan SmallCap Dividend	DGS Emg Mkts SmallCap	DIM Intl MidCap Dividend	DLS Intl SmallCap Dividend	DNL Global ex-US Growth	DOL Intl LargeCap Dividend
Inception date	Jun-06	Jun-06	Jul-07	Jun-06	Jun-06	Oct-07	Jun-06	Jun-06	Jun-06	Jun-06
α. %/year	**−0.66**	**0.71**	**0.59**	**−0.87**	**−0.88**	**−1.75**	**−0.51**	**−1.04**	**−3.26**	**−0.60**
Excess St. D of fund %/month	0.33	0.17	−0.08	0.03	−0.01	−0.04	0.07	0.04	0.25	0.04
St. D. of predict error %/mo	0.02	0.02	0.02	0.01	0.01	0.01	0.86	1.04	2.06	0.92
Correlation	95.6	97.1	96.7	98.2	96.9	98.2	99.0	98.4	90.8	98.7
Significance one tail	40	35	39	30	25	17	32	21	10	30
DEMSX Emg Mkts Small Cap	−27	2	11	16	17	91	25	18	47	15
DFALX Large Cap Intl	−6	22	19	4	9	6	38	−4	231	124
DFCSX Continental Small Co	75	−5	10	83	−16	12	25	32	−18	41
DFEMX Emerging Markets	123	100	117	1	7	75	10	19	38	13
DFEVX Emg Mkts Value	−84	−48	−29	−9	−17	−49	−21	−28	−54	−20
DFISX Intl Small Co	−228	32	−18	−66	52	−46	−10	7	−79	−75
DFIVX International Value	5	0	−34	−25	−7	−39	11	−1	−161	−12
DFJSX Japanese Small Co	43	3	6	3	84	4	9	20	−12	13
DFRSX Asia Pacific Small Co	111	26	−11	−2	−14	−13	−8	11	−30	2
DFUKX U.K. Small Co	26	−5	11	47	−4	15	9	15	−8	14
DISVX Intl Small Cap Value	63	−28	18	48	−10	46	11	11	146	−14
Sum of shorts	−346	−86	−91	−102	−69	−147	−38	−33	−363	−121
α 1st half. %/year	0.69	0.93	5.87	−3.82	1.76	2.04	−1.22	−1.02	5.20	−1.33
α 2nd half. %/year	0.69	1.00	0.29	5.73	−0.85	−4.07	−1.45	0.46	−2.02	0.90
α average. %/year	**0.69**	**0.97**	**3.08**	**0.96**	**0.46**	**−1.02**	**−1.34**	**−0.28**	**1.59**	**−0.22**

DOO Intl Divend ex-Financials	DTH DEFA Equity Income	DWM DEFA	DXJ Japan Dividend	EPI India Earnings	GULF Middle East Dividend	HEDJ European Hedged Equity	**WT Equal Weights**	**WT Risk Averse. 75 % DNL & 25 % DFJ**	**Avg for individual funds**	**Median for individual funds**
Jun-06	Jun-06	Jun-06	Jun-06	Feb-08	Jul-08	Nov-09	Feb-08	Jun-06	Jan-07	Jun-06
−0.81	**−1.01**	**−0.68**	**−3.06**	**−3.06**	**−6.05**	**1.90**	**−1.71**	**−2.60**	**−1.24**	**−0.87**
0.09	0.04	0.04	0.53	1.14	2.24	0.55	0.01	0.08	0.35	0.08
1.15	1.23	0.85	2.29	4.43	5.57	4.26	0.81	1.53	1.38	0.89
98.3	98.2	98.9	87.5	90.1	36.9	11.7	99.1	93.6	85.9	97.0
29	24	26	14	22	23	40	6	9	28	29
23	14	16	54	66	11	46	29	40	19	17
49	89	105	122	29	9	18	50	176	48	21
47	58	36	−11	72	−43	−32	26	−17	27	7
30	16	11	−28	−13	146	182	39	30	57	25
−40	−27	−20	−13	141	−97	−167	−26	−45	−34	−27
−47	−97	−56	−78	−171	234	417	−56	−46	5	−46
45	17	−11	−49	−78	−30	0	−21	−123	−26	−14
14	18	11	55	42	22	7	21	12	22	13
4	9	0	−30	−40	−57	−91	−1	−26	−6	−5
15	22	13	4	25	40	−107	13	−7	7	13
−41	−18	−5	74	28	−136	−172	26	107	−20	3
−127	−142	−92	−209	−302	−362	−569	−104	−265	−212	−134
−0.95	−2.56	−1.19	−2.76	−13.03	−19.22	−8.62	−2.45	4.37	−2.31	−1.19
0.69	1.96	0.58	−3.36	−3.03	−3.03	12.80	0.32	−1.70	0.43	0.46
−0.13	**−0.30**	**−0.31**	**−3.06**	**−8.03**	**−11.13**	**2.09**	**−1.07**	**1.34**	**−0.94**	**−0.37**

- α is the annual average continuously compounded excess return of the fundamental portfolio over that of the traditional index clone over the entire period, continuously compounded, and expressed in percentage points per year.
- **Significance of α**. In all of our exhibits "significance" is the significance of α on a one tail test. It denotes the probability that the sign of α for our sample is different from the sign of the α for the universe from which the sample is drawn. For example, from the top line of Exhibit 15.1. the DFA fund with ticker DFCCX outreturned its clone traditional portfolio with no shorts permitted by α equal to 0.14 % per year, continuously compounded, and the significance of 33 % tells that the probability is 33 % that α for the universe is negative rather than positive. The statistical significance test is done using the Paired 2-Sample *t*-Test from the Microsoft Excel Data Analysis Package, using continuously compounded monthly returns.
- **Correlation** is the correlation of the continuously compounded monthly returns for the fundamental fund and its clone.
- **Excess standard deviation of fund** is the excess volatility of the fundamental portfolio compared to that of the traditional portfolio.
- **Standard deviation of prediction error** measures the standard deviation of the return differentials, not continuously compounded. During modeling, we constrained this value to be minimized through Solver.
- **The numbers near the bottom boxes are** the weights in percent given to each DFA traditional index fund to make the portfolio that mimics the compared fundamental portfolio. In every column, they add up to 100.
- **Sum of shorts** is the share of the mimic portfolio that is held short, expressed as a percent. It is zero when shorts are prohibited.
- **α 1st half and α 2nd half** are the αs for the two half periods, and **α average** is the average of these two αs. Dividing the period allows closer tracking of the style changes.

15.9 DFA Individual Funds

We use both traditional index funds and fundamental index funds from Dimensional Fund Advisors (DFA). The DFA traditional index funds constitute the clone portfolio to compare with fundamental funds. The overall investment strategy of DFA on international investment can be found at the DFA homepage (2014). More details of each fund are described in DFA's most recent prospectus (2014).

DFA's fundamental index funds are called DFA core and DFA vector funds. The weights of stocks in their portfolios are determined by fundamentals and market capitalization. Among fundamentals, growth and value are assessed by factors, such as price-to-cash flow or price-to-earnings ratios (DFA Prospectus (2014)). Six DFA fundamental index funds were selected. They are listed, proceeded by their tickers, in Exhibit 15.1. for long positions only in the clone portfolios and Exhibit 15.2. for shorts permitted.

DFA funds can only be purchased through an adviser, who charges a fee. For a list of advisors see the Retire Early Home Page (2007). Drawing from that page, for example, Asset Builder (2013) charges 0.45 % of assets annually for portfolios between $50 thousand and $250 thousand dollars and 0.30 % of assets for portfolios between $1 million and $4 million. In addition there are custodial and transactions fees. Thus, since we are using DFA traditional funds as benchmarks, these fees increase the attractiveness of RAFI and WisdomTree, beyond those presented here.

15.10 RAFI Individual Funds

RAFI funds are the PowerShares FTSE RAFI portfolios. These portfolios incorporate four fundamental factors—dividends, cash flow, sales, and book equity value—to determine each fund's weights. We examine four RAFI fundamental index funds. They are presented in Exhibits 15.3 and 15.4.

The major outlier with a low correlation and a no-shorts $\boldsymbol{\alpha}$ of 2.87 %/year is PAF (FTSE RAFI Asia Pacific ex-Japan. In November 2011 it invested 42.1 % of its assets in Australia and 36.2 % in the Republic of Korea. In April 2014 its top ten holdings were also from those two countries. There were no comparable DFA traditional index funds, so the huge **$\boldsymbol{\alpha}$ averages** of close to 4 %/year may reflect the outperformance of the average stocks in those countries rather than better than average stock selection from those two countries. The RAFI funds are described in Powershares (2014) and ResearchAffiliates (2014).

15.11 WisdomTree Individual Funds

As described on its website, WisdomTree considers fundamentals such as dividends and earnings to reflect a company's appeal (WisdomTree (2014)). So the funds focus either on dividends or earnings. Seventeen fundamental index funds were selected from WisdomTree. They are presented in Exhibits 15.5 and 15.6.

15.12 The Individual Fundamental Index Funds

For the individual funds 17/24$^{\text{ths}}$ (71 %) of the DFA $\boldsymbol{\alpha}$s are negative, 12/16$^{\text{ths}}$ (75 %) of the RAFI $\boldsymbol{\alpha}$s are positive and 44/68$^{\text{ths}}$ (65 %) of the WT $\boldsymbol{\alpha}$s are negative. The average individual-fund $\boldsymbol{\alpha}$s have the same sign pattern.

Thus, focusing on average and median αs, DFA and WisdomTree fundamental index funds had lower returns and RAFI fundamental funds had higher returns than DFA's corresponding traditional index fund portfolios. The weak significance levels mean that these calculations provide some but not a lot of guidance for what to expect from these families in the future.

15.13 Fundamental Index Portfolios

Exhibits 15.1, 15.2, 15.3, 15.4, 15.5, and 15.6 also describe the simulations for equally weighted portfolios, and for the portfolios with minimum risk. Minimum risk is minimum standard deviation of return over the sample period. Portfolio weights are constant. Both the fundamental and the clone portfolios are rebalanced monthly. The minimum risk portfolios are

- For DFA 100 % in the International Sustainability Core I fund.
- For RAFI 94.3 % in the FTSE RAFI Developed Market ex-US Small-Mid fund and 5.7 % in the FTSE RAFI Developed Markets ex-US fund.
- For WisdomTree we constrained the maximum value of the portfolio in Japanese equities to be 25 %. The constrained risk-averse portfolio is 75 % in the Global ex-US Growth fund and 25 % in the Japan SmallCap Dividend fund.

15.14 DFA Aggregates

Calculations for the aggregates are in the right four columns in Exhibits 15.1, 15.2, 15.3, 15.4, 15.5, and 15.6. For each fund family, they consist of the calculations for the two portfolios and the average and median values for individual funds.

Although August 2008 marks the latest inception of a DFA fundamental index fund, the data was obtainable only from September 2008. We use monthly returns data, which are calculated at the end of each month. For the DFA equally weighted fundamental index portfolio we have observations: from September 2008 to June 2014, and for the risk-averse portfolio we have observations from April 2008 to June 2014.

From Exhibits 15.1 and 15.2, 14 of the 16 aggregate DFA αs are negative. Our benchmark is traditional funds issued by the same company, so the smallness of these numbers for the equally weighted portfolio is not surprising, nor is the weak significance. The large negative αs for the risk adverse portfolio are a surprise.

Minimum, median, average and maximum values for the 16 portfolio alphas are reported in Exhibit 15.7. For DFA the average of the aggregate αs is −0.33 %/year. Our take-away is that the DFA fundamental funds offer similar returns to the traditional DFA funds. This is also reflected in the very high correlations of returns: all over 99.9 %. The poor level of significance means that we cannot say with any confidence that the fundamental funds are inferior to traditional funds, but we can say that no evidence is offered for the superiority of DFA's fundamental funds over DFA's traditional index funds. The DFA simulations show that an investor who is meticulous about rebalancing would have done slightly better (aside from rebalancing costs) by holding traditional funds.

For the equal weight no-short fundamental portfolio, the expense ratio in 2010 is 0.04 % higher than for its clone portfolio, and the α is a comparable −0.06 %.

Exhibit 15.7 Aggregate alphas: RAFI outreturns

	αs in %/year					Significance	Excess SD
	Minimum	Median	Average	Maximum	fraction positive	for avg %	for avg %/month
DFA	−0.74	−0.29	−0.33	0.02	2/16	20	0.019
RAFI	0.04	0.92	1.08	2.65	All 16	35	−0.04
WisdomTree	−2.60	−0.86	−0.74	1.34	2/16	25	0.019

Last two columns exclude clone portfolios that change half way through the period

15.15 RAFI Aggregates

The evaluation of RAFI portfolios starts from October 2007, when all the RAFI fundamental funds were active and like the others carries through June 2014.

From Exhibits 15.3, 15.4, and 15.7 all of the aggregate RAFI αs are positive, ranging from 0.04 %/year to +2.65 % per year with median and average values of 0.92 %/year and 1.08 % per year respectively. The correlations for all the aggregates are all at least 96.8 %. The average excess standard deviation of return is slightly negative. Thus, the RAFI funds provided excess returns with slightly lower risk than their traditional DFA clones.

15.16 WisdomTree Aggregates

The evaluation of WisdomTree equal-weight portfolio starts from March 2008, when all the WisdomTree fundamental funds were active, except HEDJ and GULF which we excluded from the equal-weight portfolios and possible risk-averse portfolios.

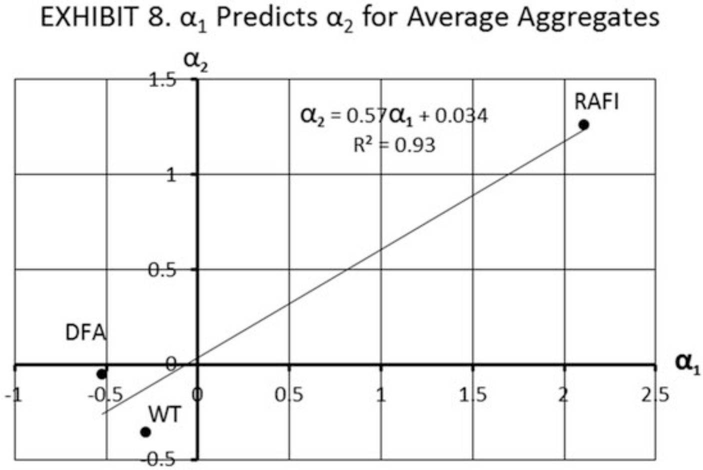
From Exhibits 15.5, 15.6 and 15.7, 14/16$^{\text{ths}}$ of the WisdomTree aggregate αs are negative, ranging from −2.60 %/year to 1.34 %/year with median and average values of −0.86 %/year and −0.74 %/year respectively. All of the correlations for the aggregates are at least 81.9 %. Six out of eight of the excess standard deviations of the aggregates are positive, with an average value of 0.019 %/month.

15.17 Does the First Half Period α Predict the Second Half Period α?

Does the α for a fundamental mutual fund in the first half of the period predict the α for the second period? Each fund family has four aggregates. This gives us eight first half αs and eight second half αs per family. We take averages for the first and second halves of the period: α_1 and α_2. These are graphed in Exhibit 15.8. The graph shows that the regression indicates a one percentage point increase in the first period α predicts a 0.57 percentage point increase in the second period α. The R^2 for the relationship is 0.93. Thus for this particular data set the relationship between the αs is positive, but the variance in the second period is smaller than that in the first period.

Our readers note the much smaller range for alphas in the second period than in the first. They write “Our guess would be that this is because the global financial crisis is in the first half of the sample, and the second half is much less turbulent by comparison. If that is the reason for the change in dispersion of alpha from the first half of the sample to the second half, then that would explain the low slope.”

Exhibit 15.8 α_1 predicts α_2 for average aggregates



15.18 Are the αS Explained by Different Sector Returns?

Bill Bernstein suggested that some of our αs might be explained by the different sectors that different funds invest in. For example, the WT equally weighted portfolio invested roughly 5 % more of its portfolio over the period in financial services and roughly 5 % less in technology than did its DFA clone (shorts permitted).

From the January 2012 Morningstar Principia disk, we drew data on the average share of investment in 11 sectors for the RAFI and WT equally weighted portfolios and their DFA clones. For convenience, we used the 3-year average ending in September 30, 2011, rather than the precise average over the entire period. We calculated the fundamental fund share in each sector minus that for the DFA clone: call it the share difference. Fidelity has sector funds that correspond to each of the 11 sectors. They are labeled Fidelity Select funds. We used Yahoo finance to gather the geometric average, continuously compounded rate of return for each Fidelity fund over the life of the equally weighted portfolio: call it return. We assumed that these Fidelity returns are good proxies for the returns to the corresponding sectors worldwide. We multiplied share difference by return for each sector and summed over the entire 11 sectors to get: the part of α that is explained by share difference: call it αShare.

The α for the equally weighted WT portfolio versus the short DFA clone through December 2012 is minus 2.15 %/year. The corresponding αShare is minus 1.22 %/year. Thus over half of the α is explained by sector choice.

The α for the equally weighted WT portfolio versus the long DFA clone, also ending in December 2012, is minus 1.35 %/year. The corresponding αShare is minus 0.40 %/year. Thus 30 % of the α is explained by sector choice.

For the RAFI equal-weight portfolio compared with the no-shorts clone the α calculated through December 2012 is +0.91 %/year. The αShare is +0.075/year. Thus less than 10 % of its α is explained by sectoral choice.

The Fidelity Select funds are primarily composed of US stocks, so they are not a great proxy for the performance of the international sectors, so it is likely that we have under-estimated the impacts of sectoral differences in our clone portfolios.

15.19 Conclusion

We have analyzed portfolio returns for the international fundamental index funds from DFA, RAFI, and WisdomTree, relative to the corresponding traditional index funds from DFA. These average αs for the 16 calculations for the four aggregates for each of the three fund families are −0.33, 1.08, and −0.74 % per year. The corresponding average significance levels on a one tailed test are 20, 35, and 25 %. Thus DFA slightly under-returned, RAFI out-returned, and WT under-returned, but the αs are not significant at standard levels.

The average standard deviations of the returns of each of the aggregates for the fundamental index funds for DFA and WisdomTree were higher than the clone portfolios (both by 0.019 %/month), but for RAFI it was lower (−0.040 %/month). In the Tower Yang (2013) paper the corresponding averages were all higher.

The corresponding average αs for US fundamental index fund portfolios relative to Vanguard US index funds from Tower and Yang (2013) are −1.43, +2.57, and 2.21 %/year. Thus both domestic and international DFA under-returned, while both domestic and international RAFI out-returned. The WisdomTree US portfolio out-returned but the WisdomTree international aggregates under-returned. Putting the conclusions of both studies together: three cheers for fundamental indexation and three cheers for traditional indexation. Thus on average, based on these calculations, we cannot rank the two methods of indexing.

These calculations are for short periods, so we hesitate to make too much of them. Still, we do not know a better way to form priors for what future excess returns of fundamental index funds are likely to be. We hope others using additional data and alternative approaches will shine a brighter light on the question.

Research Affiliates is a definite winner from this study. This rash conclusion, however, should be tested further with longer time span in the future. The study presented in the chapter, with its mixed results, can be summarized by the same quotation from Bernstein (2006) with which Tower and Yang concluded their study: "The prospective shareholder needs to consider not only the selection paradigm used, but just who is executing it."

Acknowledgments This chapter is an update and revision of Lim's honors thesis, which was supervised by Tower. Her thesis is available from her upon request. Bill Bernstein, Henry Blinder, James Dean, John Guerard, Mel Lindauer, and Steve Thorpe provided helpful suggestions. Rob Arnott, Noah Beck, and Vitali Kalesnik kindly commented on a near final version of the chapter. The latter are the readers we have quoted in several instances.

References

Arnott, R. D., Hsu, J. C., & West, J. M. (2008). *The fundamental index: A better way to invest*. Hoboken, NJ: Wiley.

Arnott, R. D., & Fabozzi, F. J. (1992). *Active Asset Allocation: State of the Art Portfolio Policies, Strategies, and Tactics*. New York City, NY: McGraw-Hill.

Asset Builder (2013). Assetbuilder Fees. Retrieved September 14, 2014 from http://assetbuilder.com/investing/assetbuilder_fees

Anderson Jenny. (August 15, 2006). Real Money Rides on Bet on Basics of Investing. *The New York Times*. http://www.nytimes.com/2006/08/15/business/15place.html?ex=1313294400&en=d43f3343f3a02979&ei=5090&partner=rssuserland&emc=rss

Bernstein, W. J. (2006). Fundamental Indexing and the Three-Factor Model. *Efficient frontier*, http://www.efficientfrontier.com/ef/0adhoc/fi.htm.

DFA. DFA home page Non-US Equity section. http://www.dfaus.com/strategies/non-us-equities.html. Retrieved September 14, 2014.

DFA Prospectus. (February 28, 2014) http://www.dfaus.com/pdf/prospectuses/idg_equity_i_pros.pdf

Lucas, Lori, and Riepe, Mark W. (1966) The Role of Returns-Based Style Analysis: Understanding, Implementing, and Interpreting the Technique. http://corporate.morningstar.com/ib/documents/MethodologyDocuments/IBBAssociates/ReturnsBasedAnalysis.pdf

Powershares. Fundamental Index Methodology. *Invesco PowerShares Capital Management LLC*. Retrieved September 27, 2014 from http://invescopowershares.com/pdf/P-FUNDAMENTALS-IVG-1.pdf

Research Affiliates. RAFI Fundamental Index. Retrieved September 27, 2014 from http://www.researchaffiliates.com/Our%20Ideas/Solutions/RAFI/RAFI%20Equity/Pages/Home.aspx

Retireearlyhomepage. (2007). Low-cost, fixed-fee DFA investment advisors. Posted July 1, 2007. Retireearlyhomepage.com/dfaadv.html

Sharpe, W. F. (Winter 1992). Asset Allocation: Management Style and Performance Measurement. *Journal of Portfolio Management, 18*(2), 7–19.

Tower, E. (2009). Classic and enhanced index funds: Performance and Issues. In J. A. Haslem (Ed.), *Mutual funds: Portfolio structures, analysis, management, and stewardship* (pp. 237–264). Hoboken, NJ: Wiley.

Tower, Edward, and Yang, Chao. (Summer 2013) Is Traditional Indexation Passé? Vanguard versus the enhanced indexers, DFA, RAFI, and WisdomTree. Factor Based Portfolio Construction. *Journal of Index Investing,* 4(1), 1–12.

WisdomTree. Retrieved September 14, 2014 from http://www.wisdomtree.com/etfs/

Chapter 16
Forecasting Implied Volatilities for Options on Index Futures: Time-Series and Cross-Sectional Analysis versus Constant Elasticity of Variance (CEV) Model

Tzu Tai and Cheng Few Lee

16.1 Introduction

Forecasting volatility is crucial to risk management and financial decision for future uncertainty. Previous studies have found that the volatility changes are predictable (Engle, 1982; Pagan & Schwert, 1990; Harvey & Whaley, 1991, 1992a, 1992b; Day & Lewis, 1992; Fleming, 1998). In perfectly frictionless and rational markets, options and their underlying assets should simultaneously and properly change prices to reflect new information. Otherwise, costless arbitrage profits would happen in portfolios combined by options and their underlying assets. However, prices in security and option markets may differently and inconsistently change to respond to news because transaction costs vary cross financial markets (Phillips & Smith, 1980). Based on trading cost hypothesis, the market with the lowest trading costs would quickly respond to new information. The price changes of options on index and options on index futures lead price changes in the index stocks because trading costs of index option markets are lower than the cost of trading an equivalent stock portfolio (Fleming, Ostdiek, & Whaley, 1996). Therefore, the dynamic behavior of market volatility can be captured by forecasting implied volatilities in index option markets (Dumas, Fleming, & Whaley, 1998; Harvey & Whaley, 1992a).

In this chapter, we use option prices instead of relying on the past behavior of asset prices to infer volatility expectations of underlying assets. The derivation and use of the implied volatility (called IV hereafter) for an option as originated by Latane and Rendleman (1976) has become a widely used methodology for variance

T. Tai (✉)
Mezocliq, LLC, New York, NY 10019-5905, USA
e-mail: tzutai30@gmail.com

C.F. Lee
Rutgers University, New Brunswick, NJ 08901, USA
e-mail: lee@business.rutgers.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_16

estimation. The IV derived from option prices depends on the assumptions of option valuation formula. For example, IV in Black-Scholes-Merton option pricing model (called BSM hereafter) tends to differ across exercise price and times to maturity, which violates the assumption of the constant volatility of underlying asset in model. The fact that there are as many BSM IV estimates for an underlying asset as there are options on it, as well as the observable nonconstant nature, has attracted considerable attention from practitioner and theoretician alike.

For the academician, previous studies have been proposed to capture the characteristics of implied volatility by either using statistical models or stochastic diffusion process approaches. Statistical models such as autoregressive conditional heteroskedasticity (ARCH) models (Engle, 1982) and GARCH model (Day & Lewis, 1992) have been used to capture time-series nature of IV dynamic behavior. On the other hand, stochastic process models such as constant-elasticity-of-variance (CEV) model (Cox, 1975; Cox & Ross, 1976; Beckers, 1980; Chen & Lee, 1993; DelBaen & Sirakawa, 2002; Emanuel & MacBeth, 1982; MacBeth & Merville, 1980; Hsu et al., 2008; Schroder, 1989; Singh & Ahmad, 2011; Pun & Wong, 2013; Larguinho et al., 2013) and stochastic volatility models (Hull & White, 1987; Heston, 1993; Scott, 1997; Lewis, 2000; Lee, 2001; Jones, 2003; Medvedev & Scaillet, 2007) incorporate the interactive behaviors of an asset and its volatilities in option pricing model. From the practitioner's point of view, the implementation and computational costs are the principal criteria of selecting option pricing models to estimate IV. Therefore, we use cross-sectional time-series regression and CEV model to forecast IV with less computational costs.

The two alternative approaches used in this chapter give different perspective of estimating IV. The cross-sectional time-series analysis focuses on the dynamic behavior of volatility in each option contracts. The predicted IV obtained from the time-series model is the estimated conditional volatility based on the information of IV extracted from BSM. Although the estimated IVs in a time-series model vary across option contracts, this kind of model can seize the specification of time-vary characteristic that links ex post volatility to ex ante volatility for each option contract. In addition, cross-sectional analysis can capture other trading behaviors such as week effect and in/out of the money effect. On the other hand, CEV model generalizes implied volatility surface as a function of asset price. It can reduce more computational and implementation costs rather than the complex models such as jump-diffusion stochastic volatility models because there is only one more variable compared with BSM. Although the constant estimated IV for each trading day may cause low forecast power of whole option contacts, it is more reasonable that the IVs of underlying assets are independent of different strike prices and times to expiration.

The focuses of this chapter are (1) to improve the ability to forecast the IV by cross-sectional time-series analysis and CEV model, (2) to explain the significance of variables in each approaches, (3) compare prediction power of these two alternative methods, and (4) test market efficiency by building an arbitrage trading strategy. If volatility changes are predictable by using cross-sectional time-series analysis and CEV model, the prediction power of these two methods can draw specific implications as to how BSM might be misspecified. If the abnormal returns are impossible in a trading strategy which takes transaction costs into account, we would claim that option markets are efficient.

The structure of this chapter is as follows. Section 16.2 reviews previous option pricing models and related empirical works concerning the viability and use of these models. The data and methodology are described in Sect. 16.3. Section 16.4 shows the empirical analysis and devise the trading and hedging strategies to determine if arbitrage profit can be obtained. Finally, in Sect. 16.5, the implications of the results are summarized from both an academic and practitioner view.

16.2 Literature Review

The amount of option pricing research is substantial. This section briefly surveys the major studies which form the impetus for this research effect. Then we introduce previous literature using time-series analysis as an alternative approach to forecast implied volatilities.

16.2.1 Black-Scholes-Merton Option Pricing Model (BSM) and CEV Model

Option pricing is a central issue in the derivatives literature. After the seminal papers by Black and Scholes (1973) and Merton (1973), there has been an explosion in option pricing models developed over the last few decades (Black, 1975; Brenner et al., 1985; Chance, 1986; Ramaswamy & Sundaresan, 1985; Wolf, 1982; Hull, 2011). BSM formula for a European call option on a stock with dividend yield rate, q, is:

$$C_t = S_t e^{-q\tau} N(d_1) - \mathrm{K}e^{-r\tau} N(d_2) \tag{16.1}$$

where $d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - q + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}$, $d_2 = d_1 - \sigma\sqrt{\tau}$, $N(\cdot)$ is the cumulative probability distribution function for a standardized normal distribution, τ is time to maturity, r is risk free rate, S_0 is current underlying stock price, K is exercise price, σ^2 is the variance of stock returns, C_t is the theoretical BSM option price at time t.

Black's (1976) model for pricing futures call options is used in this study. His model is:

$$\begin{aligned} C_t^F &= e^{-r\tau}\left[F_t N(d_1) - \mathrm{KN}(d_2)\right] \\ d_1 &= \left[\ln(F_t/K) + \left(\sigma_f^2/2\right)\tau\right]/\sigma_f\sqrt{\tau} \\ d_2 &= d_1 - \sigma_f\sqrt{\tau} \end{aligned} \tag{16.2}$$

where C_t^F is the model price for a call option on future at time t, F_t is the underlying futures price at time t, K is the exercise price of the call option, τ is the option's remaining time to maturity in terms of a year, r is the continuous annualized risk-free rate, ${\sigma_f}^2$ is the instantaneous variance of returns of the underlying futures contract over the remaining life of the option.

Although it is well known that the BSM model exhibits biases in its pricing of deep-in and out-of-the-money options and those with a very short or very long term to maturity, the direction of the bias has not been consistent across studies. Black (1975) found that the BSM model systematically over-priced options which were deep-in-the-money and underpriced those being deep-out-of-the-money. However, MacBeth and Merville (1979) reported an exactly opposite type of systematic bias. To make matters even more imprecise, Merton (1976) notes that practitioners often claim that the BSM underprices both deep-in and out-of-the-money options. In regards to time to maturity, it is generally maintained that the BSM underprices short-maturity and overprices long-maturity options. But again, the evidence contains discrepancies, particularly when the bias relative to both exercise price and maturity are considered. All these authors conclude that, to some degree, the pricing bias is related to the volatility parameter which is typically observed not to be proportionally constant over time. Jarrow and Rudd (1982) focus on the potential effects from distributional misspecification of the underlying return-generating process. Thus, their model takes into account pricing biases which might arise due to differences between the second, third and fourth moments of the assumed and "true" distributions.

Previous studies have shown that the constant volatility assumption is inappropriate, and the evidence of our empirical results presents as well. Several more generalized models have been proposed to overcome the BSM restriction on the volatility parameter. Cox (1975) and Cox and Ross (1976) developed the "constant elasticity of variance (CEV) model" which incorporates an observed market phenomenon that the underlying asset variance tends to fall as the asset price increases (and vice versa). The advantage of CEV model is that it can describe the interrelationship between stock prices and its volatility. The constant elasticity of variance (CEV) model for a stock price, S, can be represented as follows:

$$dS = (r - q)\,Sdt + \delta S^{\alpha} dZ \tag{16.3}$$

where r is the risk-free rate, q is the dividend yield, dZ is a Wiener process, δ is a volatility parameter, and α is a positive constant. The relationship between the instantaneous volatility of the asset return, $\sigma(S,t)$, and parameters in CEV model can be represented as:

$$\sigma\,(S,t) = \delta S^{\alpha - 1} \tag{16.4}$$

When $\alpha = 1$, the CEV model is the geometric Brownian motion model we have been using up to now. When $\alpha < 1$, the volatility increases as the stock price decreases. This creates a probability distribution similar to that observed for equities with a heavy left tail and a less heavy right tail. When $\alpha > 1$, the volatility increases as the stock price increases, giving a probability distribution with a heavy right tail and a less left tail. This corresponds to a volatility smile where the implied volatility is an increasing function of the strike price. This type of volatility smile is sometimes observed for options on futures.

The formula for pricing a European call option in CEV model is:

$$C_t = \begin{cases} S_t e^{-q\tau}\left[1-\chi^2(a,\ b+2,\ c)\right] - \mathrm{K}e^{-r\tau}\chi^2(c,\ b,\ a) & \text{when } \alpha < 1 \\ S_t e^{-q\tau}\left[1-\chi^2(c,-b,\ a)\right] - \mathrm{K}e^{-r\tau}\chi^2(a,\ 2-b,\ c) & \text{when } \alpha > 1 \end{cases} \tag{16.5}$$

where $a = \frac{\left[\mathrm{K}e^{-(r-q)\tau}\right]^{2(1-\alpha)}}{(1-\alpha)^2\upsilon}$, $b = \frac{1}{1-\alpha}$, $c = \frac{S_t{}^{2(1-\alpha)}}{(1-\alpha)^2\upsilon}$, $\upsilon = \frac{\delta^2}{2(r-q)(\alpha-1)}\left[e^{2(r-q)(\alpha-1)\tau}-1\right]$, and $\chi^2(z,\ k,\ v)$ is the cumulative probability that a variable with a noncentral χ^2 distribution[1] with noncentrality parameter v and k degrees of freedom is less than z. Hsu, Lin and Lee (2008) provided the detailed derivation of approximative formula for CEV model. Based on the approximated formula, CEV model can reduce computational and implementation costs rather than the complex models such as jump-diffusion stochastic volatility model. Therefore, CVE model with one more parameter than BSM can be a better choice to improve the performance of predicting implied volatilities of index options (Singh & Ahmad, 2011).

Beckers (1980) investigate the relationship between the stock price and its variance of returns by using an approximative closed-form formulas for CEV model based on two special cases of the constant elasticity class ($\alpha = 1$ or 0). Based on the significant relationship between the stock price and its volatility in the empirical results, Beckers (1980) claimed that CEV model in terms of noncentral Chi-square distribution performs better than BC model in terms of log-normal distribution in description of stock price behavior. MacBeth and Merville (1980) is the first paper to empirically test the performance of CEV model. Their empirical results show the negative relationship between stock prices and its volatility of returns; that is, the elasticity class is less than 2 (i.e., $\alpha < 2$). Jackwerth and Rubinstein (2001) and Lee, Wu, and Chen (2004) used S&P 500 index options to do empirical work and found that CEV model performed well because it took account the negative correlation between the index level and volatility into model assumption. Pun and Wong (2013) combine asymptotics approach with CEV model to price American options. Larguinho et al. (2013) compute Greek letters under CEV model to measure different dimension to the risk in option positions and investigate leverage effects in option markets. Tsai (2014) applied CEV model to portfolio hedge strategy and found CEV model can reduce replication error of barrier call options.

[1]The calculation process of $\chi^2(z,\ k,\ v)$ value can be referred to Ding (1992). The complementary noncentral chi-square distribution function can be expressed as an infinite double sum of gamma function, which can be referred to Benton and Krishnamoorthy (2003).

Merton (1976) derived a model based on a jump-diffusion process for the underlying security that allows for discontinuous jumps in price due to unexpected information flows. Geske (1979) derived a compound-option formula which considers the firm's equity to be an option underlying the exchange traded option. An interesting feature of Geske's model is that by incorporating the effects of a firm's leverage on its option the model allows for a nonconstant variance. Alternative option pricing models to describe nonconstant volatility is stochastic volatility models which consider the volatility of the stock as a separate stochastic factor (Scott, 1987; Wiggins, 1987; Stein & Stein, 1991; Heston, 1993; Lewis, 2000; Lee, 2001; Jones, 2003; Medvedev & Scaillet, 2007). Heston (1993) assumes the dynamics of instantaneous variance, V, as a stochastic process:

$$dS = \mu S dt + \sqrt{V} S dZ_1 \tag{16.6}$$

$$dV = (\alpha + \beta V)\, dt + \sigma \sqrt{V} dZ_2 \tag{16.7}$$

where dZ_1 and dZ_2 are Wiener processes with correlation ρ. For the complex implied volatility model without closed-form solutions, advanced techniques such as partial differential equations (PDEs) or Monte Carlo simulation are used to estimate the approximation of implied volatility under non-tractable models. Lewis (2000) and Lee (2001) estimate implied volatility under stochastic volatility model without jumps. Jones (2003) extends the Heston model and proposes a more general stochastic volatility models in the CEV class as follows:

$$dS = \mu S dt + \sqrt{V} S dZ_1 \tag{16.8}$$

$$dV = (\alpha + \beta V)\, dt + \sigma_1 V^{\gamma_1} dZ_1 + \sigma_2 V^{\gamma_2} dZ_2 \tag{16.9}$$

where dZ_1 and dZ_2 are independent Wiener processes under the risk-neutral probability measure. The model setting in Jones (2003) allows the correlation of the price and variance processes to depend on the level of instantaneous variance. Recently, Medvedev and Scaillet (2007) deal with a two-factor jump-diffusion stochastic volatility model where there is a jump term in stock price and volatility follows another stochastic process related to stock price's Brownian motion term with constant correlation ρ. Medvedev and Scaillet (2007) empirical results advocate the necessary of introducing jumps in stock price process. They found that jumps are significant in returns. The evidence also supports the specification of the stochastic volatility in CEV model (Jones, 2003; Heston, 1993).

The optimal selection of an option pricing model should be based on a trade-off between its flexibility and its analytical tractability. The more complicated model it is, the less applicable implementation the model has. Although jump-diffusion stochastic volatility models can general volatility surface as a deterministic function of exercise price and time, the computational costs such as parameter calibration or model implementation are high. Chen, Lee and Lee (2009) indicated that CEV model should be better candidate rather than other complex jump-diffusion stochastic volatility models because of fast computational speed and less implementation costs. Therefore, we decide to use CEV model for forecasting implied volatilities in our empirical study.

16.2.2 Time-Varying Volatility and Time-Series Analysis

Several studies have attempted to improve the estimation of the volatility term required by the BSM and Black models. Harvey and Whaley (1992a, 1992b) stated that market volatility changes are predictable by forecasting the volatility implied in index options. Their findings are consistent with the trading cost hypothesis that the index futures and option price changes lead price changes in the stock market (Stephan & Whaley, 1990; Fleming, Ostdiek, & Whaley, 1996). Therefore, we can employ the predicted IV to do hedge strategy and risk management.

All the studies involving IV estimation point out to one degree or another that for any day, the individual IV's for all the options on a particular asset (stock or futures contract) will all be different, and will change over time. Yet as MacBeth and Merville (1979) aptly note, different exercise prices should not imply differing IV's since the IV pertains to the underlying asset itself and not the exercise price. In what might be considered a preliminary basis for this study, MacBeth and Merville (1979) relate systematic pricing differences between market and BSM option prices to the systematic differences that occur among individual IV's relative to exercise price and time to maturity.

Since Latana and Rendleman's (1976) development of the IV concept, numerous researchers have studied different weighting schemes in calculating the IV. The majority of studies, including Schmalensee and Trippi (1978) and Chiras and Manaster (1978), devise weighting schemes which aim at deriving a single weighted IV from among all individual IV's for input into the BSM model. Whaley (1981; 1982) and Park and Sears (1985) utilized an OLS regression procedure to weight and segregate IV's by maturity date. The major finding of the Park and Sears (1985) study, which used option on stock index futures data, was a "time-to-maturity" effect in the pattern of the weighted IV's over time. The authors interpreted their findings as being consistent with Merton's (1973) option pricing model with stochastic interest rate. This is a portion of the IV's instability is due to the diminishing instantaneous variance of the riskless security.

Another rather foreshadowing study conducted by Brenner and Galai (1981) not only found significant divergence between the daily individual IV's and some time-series average IV, but that the distributions of the average IV's were not invariant over time. Finally, Rubenstein (1985) used individual IV's to test five alternative option pricing models versus the BSM formulation, and attempted to explain observed pricing biases. Rubenstein (1985) reported that the direction of pricing bias changed over time. This instability could be a function not only of a time-varying volatility term, but also stochastic interest rates and a changing stock market climate. Harvey and Whaley (1992a, 1992b) utilized OLS regression of the change in IV on S&P 100 index option on lagged IV, week effect dummy variables, and interest rate measures to test if IV is predictable. The significant abnormal returns obtained in Harvey and Whaley (1992a, 1992b) indicated that the market volatility is predictable time-varying variable and can be estimated by time-series analysis.

16.3 Data and Methodology

16.3.1 Data

The data for this study of individual option IV's included the use of call options on the S&P 500 index futures which are traded at the Chicago Mercantile Exchange (CME).[2] The Data is the options on S&P 500 index futures expired within January 1, 2010 to December 31, 2013. The reason for using options on S&P 500 index futures instead of S&P 500 index is to eliminate from nonsimultaneous price effects between options and its underlying assets (Harvey & Whaley, 1991). The option and future markets are closed at 3:15pm Central Time (CT), while stock market is closed at 3pm CT. Therefore, using closing option prices to estimate the volatility of underlying stock return is problematic even though the correct option pricing model is used. In addition to no nonsynchronous price issue, the underlying assets, S&P 500 index futures, do not need to be adjusted for discrete dividends. Therefore, we can reduce the pricing error in accordance with the needless dividend adjustment. According to the suggestions in Harvey and Whaley (1991, 1992a, 1992b), we select simultaneous index option prices and index future prices to do empirical analysis.

The risk free rate used in Black model and CEV model is based on 1-year Treasury Bill from Federal Reserve Bank of ST. LOUIS.[3] Daily closing price and trading volumes of options on S&P 500 index futures and its underlying asset can be obtained from Datastream.

There are two ways to select data in respect to two alternative methodologies used in this chapter. For time-series and cross-sectional analysis, we ignore transaction information and choose the futures options according to the length of trading period. The futures options expired on March, June and September in both 2010 and 2011 are selected because they have over 1 year trading date (above 252 observations) while other options only have more or less 100 observations. Studying futures option contracts with same expired months in 2010 and 2011 will allow the examination of

[2]Nowadays Chicago Mercantile Exchange (CME), Chicago Board of Trade (CBOT), New York Mercantile Exchange (NYMEX), and Commodity Exchange (COMEX) are merged and operate as designated contract markets (DCM) of the CME Group which is the world's leading and most diverse derivatives marketplace. Website of CME group: http://www.cmegroup.com/

[3]Website of Federal Reserve Bank of ST. LOUIS: http://research.stlouisfed.org/

IV characteristics and movements over time as well as the effects of different market climates.

In order to ensure reliable estimation of IV, we estimate market volatility by using multiple option transactions instead of a single contract. For comparing prediction power of Black model and CEV model, we use all futures options expired in 2010 and 2013 to generate implied volatility surface. Here we exclude the data based on the following criteria: (1) BS IV cannot be computed, (2) trading volume is lower than 10 for excluding minuscule transactions, (3) time-to-maturity is less than 10 days for avoiding liquidity-related biases, (4) quotes not satisfying the arbitrage restriction: excluding option contact if its price larger than the difference between S&P500 index future and exercise price, and (5) deep-in/out-of-money contacts where the ratio of S&P500 index future price to exercise price is either above 1.2 or below 0.8.

After arranging data based on these criteria, we still have 30,364 observations of future options which are expired within the period of 2010 to 2013. The period of option prices is from March 19, 2009 to November 5, 2013.

16.3.2 Methodology

In this section, two alternative approaches to estimate IVs are introduced. We first illustrate how to obtain BSM IV for each option contract in MATLAB. Then, based on BSM IVs, we forecast future BSM IVs for each option contract by time-series analysis and cross-sectional regression. Finally, the second method to estimate future IV is based on CEV model. To deal with moneyness- and maturity-related biases, we use the "implied-volatility matrix" to find proper parameters in CEV model. Then, the IV surface can be represented for predicting future IV in different moneyness and time-to-maturity categories.

16.3.2.1 Estimating BSM IV

This chapter can utilize financial toolbox in MATLAB to calculate the implied volatility for futures option that the code of function is as follows:

$$\text{Volatility} = \text{blsimpv}\,(\text{Price},\ \text{Strike},\ \text{Rate},\ \text{Time},\ \text{Value},\ \text{Limit},\ \text{Tolerance},\ \text{Class})$$

where the blsimpv is the function name; Price, Strike, Rate, Time, Value, Limit, Tolerance, and Class are input variables; Volatility is the annualized IV.[4] The

[4]Detailed information of the function and example of calculating the implied volatility for futures option can be found on MathWorks website: http://www.mathworks.com/help/toolbox/finance/blkimpv.html

advantages of this function are the allowance of the upper bound of implied volatility (Limit variable) and the adjustment of the implied volatility termination tolerance (Tolerance variable), in general, equal to 0.000001. The algorithm used in blsimpv function is Newton's method.

When we do the comparison of performance between CEV model and Black model, the implied volatility of Black model for each group at time t can be obtained by following steps:

1. Let $C^F_{i,n,t}$ is market price of the nth option contract in category i, $\widehat{C^F_{i,n,t}}(\sigma)$ is the model option price determined by Black model in Eq. (16.2) with the volatility parameters, σ. For nth option contract in category i at date t, the difference between market price and model option price can be described as:

$$\varepsilon^F_{i,n,t} = C^F_{i,n,t} - \widehat{C^F_{i,n,t}}(\sigma) \tag{16.10}$$

2. For each date t, we can obtain the optimal parameters in each group by solving the minimum value of absolute pricing errors (minAPE) as:

$$\text{minAPE}_{i,t} = \min_{\sigma} \sum_{n=1}^{N} \left|\varepsilon^F_{i,n,t}\right| \tag{16.11}$$

   Where N is total number of option contracts in group i at time t.

3. Using MTALAB optimization function to find optimal σ_0 in a fixed interval. The function code is as follows:

$$[\sigma_0,\ \text{fvalBls}] = \text{fminbnd}\left(\text{fun},\ x_1, x_2\right), \tag{16.12}$$

Where σ_0 is an optimal implied volatility in Black model that locally minimize function of minAPE, fvalBls is the minimum value of minAPE, fun is MATLAB function describing Eq. (16.11). The implied volatility, σ_0, is constrained in the interval between x_1 and x_2, that is, $x_1 \leq \sigma_0 \leq x_2$. The algorithm of fminbnd function is based on golden section search and parabolic interpolation.

16.3.2.2 Forecasting IV by Cross-Sectional and Time-Series Analysis

Time-Series Analysis

Box and Jenkins (1970) time-series model building techniques are used to identify, estimate, and check models describing particular generating processes. These models are of the form

$$x_t - \Phi_1 x_{t-1} - \cdots - \Phi_p x_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{16.13}$$

where x_t is an observation from a covariance stationary series meaning that

$$\lambda_\tau = \text{cov}\,(x_t, x_{t-\tau}) \tag{16.14}$$

is independent of t for all τ. The Φ and θ terms represent the autoregressive (AR) and moving average (MA) coefficients and ε_t is white noise.

A developed technique motivated by Hannan and Rissanen (1982) seems to provide a good practical basis for model selection. The process involves two stages of computation. The purpose of the first stage is to obtain estimates of the innovation errors of model. This is accomplished by running successively higher order autoregressive models and using the AIC of Akaike (1969) to determine the optimal order from among them. The innovation errors are estimated by

$$\widehat{\varepsilon}_t = x_t - \widehat{\Phi}_1 x_{t-1} - \cdots - \widehat{\Phi}_k x_{t-k} \tag{16.15}$$

where k is the optimal autoregressive order suggested by the AIC. The second stage involves fitting all different combinations of ARMA (p, q) models where, instead of using full maximum likelihood estimation, the innovation errors estimated in stage one are used as the regressors upon which the moving average parameter estimates are based. This allows use of least squares. The different ARMA (p, q) models are then compared using the AIC of Akaike (1977) and SBC of Schwarz (1978) and the appropriate model is chosen on that basis. A simulation study conducted by Ansley and Newbold (1980) has found that exact maximum likelihood estimation outperforms least squares when the series are of moderate size and moving average terms are involved. An approximation to the full maximum likelihood function has been derived by Hillmer and Tiao (1979).

In addition, alternative simple time-series methods are taken into account to compare with the forecastability indicators from optimal ARMA models. There are five alternative models to generate IV indicators which are used in cross-sectional regression model in next section. These time-series models are as follows:

1. ARMA model (ARMA): $\text{IV}_t = a_0 + \sum_{i=1}^{p} a_i \text{IV}_{t-i} + \varepsilon_t + \sum_{i=1}^{q} b_i \varepsilon_{t-i}$
2. Lag IV method (LIV): $\text{IV}_t = \text{IV}_{t-1}$
3. 5-day moving average method (MAV5): $\text{IV}_t = \frac{\sum_{i=1}^{5} \text{IV}_{t-i}}{5}$
4. 5-day exponential moving average method (EMA5): $\text{IV}_t = \frac{\sum_{i=1}^{5} 2^{i-1} \text{IV}_{t-i}}{\sum_{i=1}^{5} 2^{i-1}}$
5. Regression on lag IV (RGN): $\text{IV}_t = a_0 + a_1 \text{IV}_{t-1} + \varepsilon_t$

The optimal ARMA model is autoregressive-moving-average model with order of the autoregressive part, p, and the order of the moving average part, q where the suitable p and q are based on the goodness-of-fit indicators, AIC and SBC, and the forecastability indicators, RMSE, MAE, and MAPE. The 5-day moving average and the 5-day exponential moving average methods can be expressed as the special cases of the general AR(5) model. Lag IV and the regression on lag IV methods belong to AR(1) process.

Cross-Sectional Predictive Regression Model

A significant amount of information has been shown to exist in a time series of IV. The five alternative time-series models used to describe the generating processes of the IV series examined are all clearly preferred to random walk or "white noise" alternatives. These models do not give the final word on the subject of IV forecasting, however. There are several cross-contract effects that may exist which, if isolated properly, will provide further predictive power. To learn more about these different influences, a large cross-sectional time-series predictive regression model was formulated. The cross-sectional time-series predictive regression model is

$$y_{it} = \beta_0 + \beta_1 x_{1it-1} + \beta_2 x_{2it-1} + \cdots + \beta_{14} x_{14it-1} + \varepsilon_{it} \tag{16.16}$$

where y_{it} is IV of the i^{th} option contract at time t; x_{1it-1} is the time-series predictor of i^{th} contract for time t based on information known at time $t-1$ and one of forecasting time-series methods; x_{2it-1} is time to maturity of the i^{th} option contract at time $t-1$ which is the unit of year; x_{3it-1} is proportional in-the-money that is equal to the value of (future price at time $t-1$—strike price)/(strike price) if the value is positive, otherwise is zero; x_{4it-1} is proportion out-of-the-money that is equal to the value of (strike price—futures price at time $t-1$)/(strike price) if the value is positive, otherwise is zero; x_{5it-1} and x_{6it-1} are standard deviation of the IV based on previous 5 and 20 observations, respectively; x_{7it-1} and x_{7it-1} are skewness and kurtosis of IV distribution over the previous 20 observations, respectively; x_{9it-1} and x_{10it-1} are the standard deviations of the rate of returns of the underlying future price on previous 5 and 20 observations, respectively; x_{11it-1}, x_{12it-1}, x_{13it-1}, and x_{14it-1} are dummy variables that equals 1 if the trading date at time $t-1$ is Tuesday, Wednesday, Thursday, and Friday, respectively.

The time-to-maturity variable was included because, as was indicated by Park and Sears (1985), there tends to be a certain point close to maturity where the IV's begin to decrease. The third and fourth independent variables have been included to see if deep-in-the-money options and far-out-of-the-money options tend toward higher or lower than expected IV's. Previous studies have had conflicting answers to this important question (see Jarrow & Rudd, 1983). The next two independent variables are included to determine whether or not the standard deviations of the IVs have any positive or negative effect on the IVs themselves. The third and fourth moments of the distribution of 20 previous IV observations were also included in the regression equation to see what, if any, influence they have in determining current IV.

The two measures of the standard deviations of the rate of returns of the underlying future price are of great interest as regressors since these have traditionally been approximations of the variable used in the BSM model to determine the theoretical option price. The final four explanatory variables are weekday effect dummies which are intended to see if certain days give rise to higher IV than others. For example, certain economic announcements are regularly made on particular days of the week and this may have a weekday effect on IV. Note that only four dummy variables are needed to describe the 5 days of the week in order to avoid perfect multi-collinearity with the constant term.

16.3.2.3 Forecasting IV by CEV Model

To deal with moneyness- and expiration- related biases in estimating BSM IV, we use the "implied-volatility matrix" to separate option contracts and estimate parameters of CEV model in each category. The option contracts are divided into nine categories by moneyness and time-to-maturity. Option contracts are classified by moneyness level as at-the-money (ATM), out-of-the-money (OTM), or in-the-money (ITM) based on the ratio of underlying asset price, S, to exercise price, K. If an option contract with S/K ratio is between 0.95 and 1.01, it belongs to ATM category. If its S/K ratio is higher (lower) than 1.01 (0.95), the option contract belongs to ITM (OTM) category. According to the large observations in ATM and OTM, we divide moneyness-level group into five levels: ratio above 1.01, ratio between 0.98 and 1.01, ratio between 0.95 and 0.98, ratio between 0.90 and 0.95, and ratio below 0.90. By expiration day, we classified option contracts into short-term (less than 30 trading days), medium-term (between 30 and 60 trading days), and long-term (more than 60 trading days).

Since for all assets the future price equals the expected future spot price in a risk-neutral measurement, the S&P 500 index futures prices have same distribution property of S&P 500 index prices. Therefore, for a call option on index futures can be given by Eq. (16.5) with S_t replaced by F_t and $q = r$ as Eq. (16.17)[5]:

$$C_t^F = \begin{cases} e^{-r\tau}\left(F_t\left[1-\chi^2(a,\ b+2,\ c)\right]-K\chi^2(c,\ b,\ a)\right) & \text{when } \alpha < 1 \\ e^{-r\tau}\left(F_t\left[1-\chi^2(c,-b,\ a)\right]-K\chi^2(a,\ 2-b,\ c)\right) & \text{when } \alpha > 1 \end{cases} \tag{16.17}$$

where

$$a = \frac{K^{2(1-\alpha)}}{(1-\alpha)^2\upsilon},\ b = \frac{1}{1-\alpha},\ c = \frac{{F_t}^{2(1-\alpha)}}{(1-\alpha)^2\upsilon},\ \ \upsilon = \delta^2\tau$$

[5]When substituting $q = r$ into $\upsilon = \frac{\delta^2}{2(r-q)(\alpha-1)}\left[e^{2(r-q)(\alpha-1)\tau}-1\right]$, we can use L'Hospital's Rule to obtain υ. Let $x = r - q$, then

$$\lim_{x\to 0}\frac{\delta^2\left[e^{2x(\alpha-1)\tau}-1\right]}{2x(\alpha-1)} = \lim_{x\to 0}\frac{\frac{\partial\delta^2\left[e^{2x(\alpha-1)\tau}-1\right]}{\partial x}}{\frac{\partial 2x(\alpha-1)}{\partial x}} = \lim_{x\to 0}\frac{(2(\alpha-1)\tau)\delta^2\left[e^{2x(\alpha-1)\tau}\right]}{2(\alpha-1)} = \lim_{x\to 0}\frac{\tau\delta^2\left[e^{2x(\alpha-1)\tau}\right]}{1} = \tau\delta^2.$$

The procedures to obtain estimated parameters of CEV model in each category of implied-volatility matrix are as follows:

1. Let $C^F_{i,n,t}$ is market price of the nth option contract in category i, $\widehat{C^F_{i,n,t}}(\delta_0, \alpha_0)$ is the model option price determined by CEV model in Eq. (16.17) with the initial value of parameters, $\delta = \delta_0$ and $\alpha = \alpha_0$. For nth option contract in category i at date t, the difference between market price and model option price can be described as:

$$\varepsilon^F_{i,n,t} = C^F_{i,n,t} - \widehat{C^F_{i,n,t}}(\delta_0, \alpha_0) \tag{16.18}$$

2. For each date t, we can obtain the optimal parameters in each group by solving the minimum value of absolute pricing errors (minAPE) as:

$$\text{minAPE}_{i,t} = \min_{\delta_0, \alpha_0} \sum_{n=1}^{N} \left| \varepsilon^F_{i,n,t} \right| \tag{16.19}$$

   Where N is total number of option contracts in group i at time t.

3. Using optimization function in MATLAB to find a minimum value of the unconstrained multivariable function. The function code is as follows:

$$[x,\ \text{fval}] = \text{fminunc}(\text{fun},\ x_0) \tag{16.20}$$

where x is the optimal parameters of CEV model, fval is the local minimum value of minAPE, fun is the specified MATLAB function of Eq. (16.19), and x_0 is the initial points of parameters obtained in step (1). The algorithm of fminunc function is based on quasi-Newton method.

16.4 Empirical Analysis

In the empirical study section, we present the forecastability of S&P 500 index option price for two alternative models: time-series and cross-sectional analysis and CEV model. First, the statistical analysis for time-series futures option prices of the contracts expired on March, June and September in both 2010 and 2011 is summarized. Then we use time-series and cross-sectional models to analyze each individual contract and compare their forecastability of IV. Finally, we estimated IV by using CEV model and compare its pricing accuracy with Black model.

Table 16.1 Distributional Statistics for Individual IV's

Option Series[a]	Mean	Std. Dev.	CV[b]	Skewness	Kurtosis	Studentized Range[c]	Observation
Call Futures Options in 2010							
MAR10 1075	0.230	0.032	0.141	2.908	14.898	10.336	251
JUN10 1050	0.263	0.050	0.191	0.987	0.943	6.729	434
JUN10 1100	0.247	0.047	0.189	0.718	−0.569	4.299	434
SEP10 1100	0.216	0.024	0.111	0.928	1.539	6.092	259
SEP10 1200	0.191	0.022	0.117	0.982	2.194	6.178	257
Call Futures Options in 2011							
MAR11 1200	0.206	0.040	0.195	5.108	36.483	10.190	384
MAR11 1250	0.188	0.027	0.145	3.739	25.527	10.636	324
MAR11 1300	0.176	0.021	0.118	1.104	4.787	8.588	384
JUN11 1325	0.165	0.016	0.095	−1.831	12.656	10.103	200
JUN11 1350	0.161	0.018	0.113	−0.228	1.856	8.653	234
SEP11 1250	0.200	0.031	0.152	2.274	6.875	7.562	248
SEP11 1300	0.185	0.024	0.131	2.279	6.861	7.399	253
SEP11 1350	0.170	0.025	0.147	2.212	5.848	6.040	470

[a]Option series contain the name and code of futures options with information of the strike price and the expired month, for example, SEP11 1350 represents that the futures call option is expired on September, 2011 with the strike price $1350 and the parentheses is the code of this futures option in Datastream

[b]CV represents the coefficient of variation that is standard deviation of option series divided by their mean value

[c]Studentized range is the difference of the maximum and minimum of the observations divided by the standard deviation of the sample

16.4.1 Distributional Qualities of IV time series

A summary of individual IV distributional statistics for S&P 500 index futures call options in 2010 and 2011 appears in Table 16.1. Comparing the mean IV's across time periods, it is quite evident that the 2011 IV's are significantly smaller. Also, the time-to-maturity effect observed by Park and Sears (1985) can be identified. The September options in 2011 possess higher mean IV's than those maturing in June and March with the same strike price.

The other statistical measures listed in Table 16.1 are the relative skewness and relative kurtosis of the IV series, along with the studentized range. Skewness measures lopsidedness in the distribution and might be considered indicative of a series of large outliers at some point in the time series of the IV's. Kurtosis measures the peakedness of the distribution relative to the normal and has been found to affect the stability of variance (Lee & Wu, 1985). The studentized range gives an overall indication as to whether the measured degrees of skewness and kurtosis have significantly deviated from the levels implied by a normality assumption for the IV series.

Using significance tests on the results of Table 16.1 in accordance with Jarque–Bera test, the 2010 and 2011 skewness and kurtosis measures indicate a higher proportion of statistical significance. We also utilize simple back-of-the-envelope test based on the studentized range to identify whether the individual IV series approximate a normal distribution. The studentized range larger than 4 in both 2010 and 2011 indicates that a normal distribution significantly understates the maximum magnitude of deviation in individual IV series.

As a final point to this brief examination of the IV skewness and kurtosis, note the statistics for MAR10 1075, MAR11 1200, and MAR11 1250 contracts. The relative size of this contract's skewness and kurtosis measures reflect the high degree of instability that its IV exhibited during the last 10 days of the contract's life. Such instability is consistent across contracts.

However, these distortions remain in the computed skewness and kurtosis measures only for these particular contracts to emphasize how a few large outliers can magnify the size of these statistics. For example, the evidence that S&P 500 future price jumped on January 18, 2010 and plunged on February 2, 2011 cause the IV of these particular contracts sharply increasing on that dates. Thus, while still of interest, any skewness and kurtosis measures must be calculated and interpreted with caution.

16.4.2 Time-Series and Cross-Sectional Analysis for IV Series

The optimal ARMA models for the IV series are based on the goodness-of-fit indicators, AIC and SBC, and the forecastability indicators, root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

Table 16.2 shows the results of the regression model in Eq. (16.16). The time-series predictor variables calculated by different forecasting models are all significant, which should come as no surprise—IV depends on past IV value. However, the fact that other regressors were found to be significant indicates that not all of the variation in IV series is explained by the past. Time-to-maturity has the predicted positive effect. The closer an option is to expiration, the lower the IV. The in-the-money effect is significantly positive; however, the out-of-the-money give mixed insignificant influence on IV series. Merton (1976) shows that large deviations from the strike price tend to bias the BSM theoretical price downward. Therefore it is logical to expect the IV of the deep-in-the-money and far-out-of-the-money contracts to be higher because the writer of these calls runs a greater risk of being stuck in his position. However, in this study, the selected IV time series calculated by BSM model cannot show the downward characteristic obviously because the longest trading data is the option contract with the strike price close to the underlying asset.

The coefficients on the standard deviation of the IV variables give the significantly positive signals based on previous 20 observations, but show the negative effect based on previous 5 observations when the short term effect is of significance.

Table 16.2 The results of the cross-sectional time-series predictive regression model

Year	2010	2011	2010	2011	2010	2011	2010	2011	2010	2011
Variable	ARMA		Previous IV		5-day moving averages		5-day exponential moving averages		Regression	
Intercept	0.0068 (3.13)***	0.0037 (1.61)	0.0146 (6.34)***	0.0260 (11.08)***	0.0049 (1.62)	−0.0067 (−2.31)**	0.0024 (0.95)	0.0016 (0.65)	0.0088 (3.70)***	0.0137 (5.51)***
Time-series predictor	0.9362 (68.26)***	0.9634 (62.36)***	0.8798 (60.80)***	0.7827 (52.38)***	0.9576 (48.08)***	1.0546 (52.08)***	0.9650 (58.85)***	0.9582 (58.33)***	0.9064 (61.01)***	0.8574 (54.03)***
Time to Maturity	0.0038 (2.43)**	−0.0013 (−1.12)	0.0068 (4.03)***	0.0048 (3.67)***	0.0010 (0.48)	−0.0054 (−3.93)***	0.0007 (0.41)	−0.0013 (−1.00)	0.0057 (3.35)***	0.0023 (1.78)*
Proportion in-the-money	0.0718 (5.81)***	0.0914 (4.83)***	0.0880 (6.56)***	0.1536 (7.38)***	0.0904 (5.72)***	0.0507 (2.34)**	0.0694 (5.01)***	0.0732 (3.67)***	0.0757 (5.63)***	0.1383 (6.74)***
Proportion out-of-the-money	−0.0088 (−1.41)	0.0143 (1.84)*	−0.0160 (−2.36)**	−0.0056 (−0.65)	−0.0004 (−0.05)	0.0204 (2.34)**	−0.0029 (−0.42)	0.0048 (0.59)	−0.0141 (−2.08)**	0.0173 (2.04)**
σ_{IV} (5 obs.)	−0.0054 (−0.10)	−0.2240 (−5.61)***	0.0366 (0.61)	−0.1059 (−2.38)**	0.0273 (0.39)	−0.4047 (−9.02)***	−0.0242 (−0.40)	−0.3028 (−7.25)***	0.0308 (0.52)	−0.1324 (−3.03)***
σ_{IV} (20 obs.)	0.2355 (3.59)***	0.1244 (2.01)**	0.3268 (4.59)***	0.3887 (5.74)***	0.2889 (3.42)***	0.1995 (2.89)***	0.2123 (2.88)***	0.2498 (3.90)***	0.3385 (4.78)***	0.4061 (6.12)***
Skewness	0.0004 (1.60)	0.0005 (1.44)	0.0005 (1.79)*	0.0006 (1.80)*	0.0001 (0.24)	0.0002 (0.68)	−0.0001 (−0.19)	0.0000 (0.10)	0.0007 (2.27)**	0.0011 (3.12)***
Kurtosis	−0.0002 (−1.78)*	0.0003 (2.15)**	−0.0001 (−0.55)	0.0001 (0.63)	−0.0001 (−0.71)	0.0006 (4.65)***	−0.0001 (−0.67)	0.0003 (2.49)**	−0.0001 (−0.93)	0.0002 (1.40)
σ_{R_future} (5 obs.)	0.0450 (0.73)	−0.0885 (−1.24)	0.0212 (0.32)	−0.0959 (−1.21)	−0.0407 (−0.52)	−0.2633 (−3.31)***	−0.0146 (−0.21)	−0.1846 (−2.48)**	0.0266 (0.40)	−0.1116 (−1.43)
σ_{R_future} (20 obs.)	0.1285 (1.46)	0.2018 (1.85)*	0.3202 (3.37)***	0.7288 (6.11)***	0.0536 (0.46)	0.1396 (1.13)	0.0433 (0.43)	0.2973 (2.61)***	0.3259 (3.44)***	0.5721 (4.85)***

(continued)

Table 16.2 (continued)

Year	2010	2011	2010	2011	2010	2011	2010	2011	2010	2011
Variable	ARMA		Previous IV		5-day moving averages		5-day exponential moving averages		Regression	
Tuesday	−0.0001 (−0.15)	0.0002 (0.23)	−0.0003 (−0.29)	0.0010 (0.99)	−0.0009 (−0.76)	−0.0004 (−0.42)	−0.0007 (−0.70)	0.0001 (0.12)	−0.0003 (−0.30)	0.0009 (0.91)
Wednesday	0.0014 (1.48)	0.0010 (1.12)	0.0015 (1.47)	0.0015 (1.53)	0.0001 (0.10)	0.0000 (−0.01)	0.0009 (0.82)	0.0007 (0.77)	0.0015 (1.47)	0.0015 (1.52)
Thursday	0.0009 (0.96)	0.0005 (0.55)	0.0010 (1.02)	0.0008 (0.82)	0.0000 (−0.02)	−0.0005 (−0.50)	0.0008 (0.77)	0.0004 (0.38)	0.0010 (1.02)	0.0008 (0.78)
Friday	0.0009 (0.99)	0.0025 (2.70)***	0.0010 (1.01)	0.0031 (3.08)***	0.0006 (0.49)	0.0016 (1.58)	0.0010 (0.92)	0.0026 (2.68)***	0.0010 (0.99)	0.0030 (3.05)***
R^2	0.9341	0.8313	0.9219	0.7913	0.8936	0.7899	0.9182	0.8163	0.9223	0.7986
Adjusted R^2	0.9334	0.8302	0.9211	0.7899	0.8926	0.7885	0.9175	0.8151	0.9215	0.7972
S_e	0.0116	0.0132	0.0127	0.0147	0.0148	0.0147	0.0130	0.0138	0.0126	0.0144
F-value	1532.75***	733.25***	1276.67***	564.08***	908.61***	559.51***	1214.92***	661.33***	1283.62***	589.98***
Chow test	2.125***		6.108***		3.763***		2.725***		2.985***	

For the forecasting methods used in Table 16.2 are as follows: ARMA is autoregressive-moving-average model with order of the autoregressive part, **p**, and the order of the moving average part, **q** as the form $\mathbf{IV_t} = \mathbf{a}_0 + \sum_{i=1}^{p} \mathbf{a_i IV_{t-i}} + \boldsymbol{\varepsilon}_t + \sum_{i=1}^{q} \mathbf{b_i \varepsilon_{t-i}}$. The suitable **p** and **q** are based on the goodness-of-fit indicators, AIC and SBC, and the forecastability indicators, RMSE, MAE, and MAPE. Previous IV is simply an AR(1) process of the form $\mathbf{IV_t} = \mathbf{IV_{t-1}}$. The 5-day moving averages can be expressed as an AR(5) model of the form $\mathbf{IV_t} = 0.2\,\mathbf{IV_{t-1}} + 0.2\,\mathbf{IV_{t-2}} + 0.2\,\mathbf{IV_{t-3}} + 0.2\,\mathbf{IV_{t-4}} + 0.2\,\mathbf{IV_{t-5}}$. The 5-day exponential moving averages is another special case of the general AR(5) model where the parameters are restricted as follows: $\mathbf{IV_t} = \frac{16}{31}\,\mathbf{IV_{t-1}} + \frac{8}{31}\,\mathbf{IV_{t-2}} + \frac{4}{31}\,\mathbf{IV_{t-3}} + \frac{2}{31}\,\mathbf{IV_{t-4}} + \frac{1}{31}\,\mathbf{IV_{t-5}}$. The Regression is the general case of AR(1) as $\mathbf{IV_t} = \mathbf{a}_0 + \mathbf{a}_1\mathbf{IV_{t-1}} + \boldsymbol{\varepsilon}_t$. Proportion in-the-money is equal to the value of (future price at time $\mathbf{t}-1$—strike price)/strike price if the value is positive, otherwise is zero. Proportion out-of-the-money is equal to the value of (strike price—future price at time $\mathbf{t}-1$)/strike price if the value is positive, otherwise is zero. σ_{IV} (5 obs.) and σ_{IV} (20 obs.) are the standard deviations of the IV based on previous 5 and 20 observations, respectively. Skewness and Kurtosis are the skewness and kurtosis of the IV distribution over the previous 20 observations, respectively. σ_{R_future} (5 obs.) and σ_{R_future} (20 obs.) are the standard deviations of the rate of returns of the underlying future price on previous 5 and 20 observations. Tuesday, Wednesday, Thursday, and Friday are the dummy variables that equal to one if the trading date of IV observation at time $\mathbf{t}-1$ is equal to that weekday and otherwise equal to zero. The value in table is represented the estimation coefficient with the t-value in parentheses

*, **, and *** in parenthesis indicate the significance at 90, 95, and 99 % level. Total observations for 2010 and 2011 are 1530 and 2098, respectively. The chow test is therefore follow $F(15, 3598)$ distribution with critical value 2.04 at 99 % level

The skewness and kurtosis terms have consistently slight effects over two sample periods even though sometimes the effects have statistical significance. Perhaps what can be said about the lower relationship between these two statistic measures and predicted IV is that the influence of the outliers bringing about the skewness and kurtosis is already captured by other independent variables such as the time-series predictor estimated by forecasting model or the standard deviation of IV series. The coefficient on the standard deviation of the rate of returns of the underlying future price only has significantly large positive effect on IV for the 20-day measure. The strong relationship to historical standard deviations of underlying assets seems that the IV series not only response to market deviation from the functional specification of the BSM model but also reflect the market assessment of the standard deviation of underlying assets.

The weekday effect dummies indicate a significantly small Friday effect where the IV are slightly higher. This may be related to the fact that certain economic announcements are made on Friday such as employment situation or lag response to the announcements made on Thursday such as money supply and jobless claims. These economic announcements will alter the market perception of asset price volatility, especially currently the situation of economics that just came through the financial crisis and is suffering from European sovereign-debt crisis. The Friday effect might also be related to option market inactivity the day before the weekend. Further study may investigate this apparent weekday effect to explain why Friday's market may be out of line with that of other days.

Whether the estimated models change significantly over time is an important question. The parameter estimates obtained for this cross-sectional time-series model seems not consistent in 2010 and 2011 sample periods. A Chow test[6] statistic indicating structural change based on five forecasting methods are obtained in Table 16.2 for the 2010 and 2011 regressions. These values exceed the table value of 2.04 for an F random variable at the 99 % level. The chow test indicated the significant change of structure in the cross-sectional time-series predictive regression model on 2010 and 2011. It would therefore be wise for the practitioner to update parameter estimates periodically even though both 2010 and 2011 sample periods are suffering from global financial crisis.

16.4.3 Ex-Post Test for Forecastability of Time-Series and Cross-Sectional Regression Models

In this section, the practical monetary value of the IV estimates versus more naive methods is tested, to determine which might be superior from a trader's point of

[6]Chow test $F_{q,n-k} = \frac{e'_* e_* / \alpha}{e' e / (n-k)}$ where $e'_* e_*$ is restricted SSE, $e'e$ is unrestricted SSE, α represents the number of restrictions, and k is number of regression coefficients estimated in unrestricted regression.

view. In addition, we hope that these results will further support the theoretical and practical superiority of using individual IV estimates versus some weighted-IV measure.

Trading rule tests in this chapter utilize seven different estimates for IV as follows: (1) a 5-day equally weighted moving average of the IV (MAV5); (2) a 5-day exponentially weighted average of the IV (EMA5); (3) a 1-day lag of the actual IV for the option (LIV); (4) 1-day ahead simple regression forecasts of the IV (RGN); (5) 1-day ahead ARMA forecasts of the IV (ARMA); (6) 1-day ahead cross-sectional time-series predictive regression forecasts of the IV based on Eq. (16.16) (CSTS); (7) a simple-constant mean of an individual IV time series for the estimated IV of that option (MEAN).

The trading rule used is simply to buy underpriced and sell overpriced options, while taking an opposite position in the underlying futures contract according to the hedge ratio computed by the estimated IV. The holdout periods for each option are 20 trading days. Here the day count convention in Black option pricing model is used actual/actual basis. Mispricing will be identified by comparing the market price for an option with the price calculated by Black option pricing model using one of the seven IV estimates. The overpriced (underpriced) options are defined as the situation that the theoretical price calculated by Black option pricing model is smaller (larger) than the market price. The trading behavior is buying (selling) the underpriced (overpriced) future option and selling (buying) S&P 500 index future for hedge. In order to magnify the mispricing as might be seen from the eyes of a trader, ten options and ten times the hedge ratio of futures are sold or bought in opposite position respectively in each transaction. Positions are closed out once the absolute value of mispricing diminishes to a predetermined minimum level equal to 0.1. If the mispricing has reversed and is of a great enough significance larger than 0.1, the trading rule is utilized again.

In order to ascribe as much realism as possible to these tests, the following market trading costs are considered. Transaction fee per transaction of \$2.3 is determined by CME group which provides CME Globex trading platform for 24-h global access to electronic markets. Total transaction fees is transaction costs of option position + transaction costs of future position:

$$\sum_{i=1}^{n} (\$2.3 \times 10) + \sum_{i=1}^{n} (\$2.3 \times 10 \times \text{hedge ratio}_{T_i})$$

where n is the total number of times a position is opened at time T_i.

Although a portion of the margin required of a trader enter into a futures position can be put up in the form of interest earning T-bills, a substantial portion required for maintaining the margin account by the clearinghouse must be strictly in cash even for a hedge or spread position. Consequently, there is a real interest cost involved, for which we will further reduce gross trading income:

$$\text{Margin Interest Costs} = \sum_{i=1}^{n} \left(\text{RMM} \times \text{NF}_{T_i} \times R_{T_i} \times \tau_i \right)$$

where RMM is required maintenance margin from CME group,[7] NF_{Ti} indicate the number of futures contracts entered into trading which is equal to ten times hedge ratio at time $T_{i,}$ R_{Ti} is the risk free rate defined as the 3-month T-bill rate used in Black option pricing model, τ_{i} is the length of futures position holding until maturity in annual terms, and n is the total number of times a futures position is entered.

Furthermore, there is little assurance that one could buy or sell these contracts and expect to receive the closing prices reported in the paper when the market reopens the next morning. To approximate such market costs the position is penalized each time a futures position is entered and existed by "one tick" equal to 0.1 index points = $25 per contract[8]:

$$\text{Futures Liquidity Costs} = \sum_{i=1}^{n} (\$50 \times \text{NF}_{T_i})$$

where $50 = 2 × $25 represented the entered and existed cost by one tick, the market value of two price ticks; NF_{Ti} is defined as the number of futures contracts entered into trading which is equal to ten times hedge ratio at time $T_{i,}$, and n is the number of times a futures position is entered. More severe liquidity and timing costs are calculated and deducted for each option transaction:

$$\text{Option Liquidity costs} = \sum_{i=1}^{n} [\$250 \times (\text{NEPA}_{T_i} + \text{NMMO}_{Ti})]$$

where $250 = 10, (number of options bought or sold) × $250 (the market value multiplier for the option premium) × 0.1 (one tick price as the correspondingly liquidity), NEPA represented the number of exercise prices in out-of-the-money options are $5 away from underlying future prices at time T_i, and NMMO represented the percentage of maturity months out. For example, a option assumed to be expired on September 2010 and this option start to be traded on February 2010, then the NMMO on June 2010 is equal to the number of month of the period between February and June divided by the number of month of the period between February and September, that is, (6–2)/(9–2) = 4/7.

The test results are summarized in Tables 16.5. We use seven alternative methods, a cross-sectional time-series regression and six time-series models, to compute tomorrow's IV for each contract. The cross-sectional time-series (CSTS) model

[7]The minimum required maintenance for S&P 500 index futures is various in different period. For example, from Jan 28, 2008 to Oct 1st, 2008, the maintenance cost is $18,000 per future contract. However, the period during Oct 1st, 2008 to Oct 17, 2008, the required maintenance is changed to $20,250. The maintenance costs are $22,500, $24,750, $22,500, and $20,000 for other periods Oct 17, 2008–Oct 30, 2008; Oct 30, 2008–Mar 20, 2009; Mar 20, 2009–Jun 2nd, 2011; and Jun 2nd, 2011 until now.

[8]The detailed contract specifications for S&P 500 futures and options on futures can be found in CME group website: http://www.cmegroup.com/trading/equity-index/files/SxP500_FC.pdf

Table 16.3 Cumulative survey of trading results for samples in holdout period

IV estimate	Gross value of all trades	Total trading costs	Net value of all trades	Number of trade made	Net profit or loss per trade
(a) 2010					
MAV5	**1,673,339**	785,469.1	887,869.8	95	9,345.997
EMA5	1,185,108	735,197.1	449,910.6	95	4,735.901
LIV	1,325,712	**405,671**	**920,041.3**	95	**9,684.645**
RGN	1,077,990	432,747.3	645,243.1	95	6,792.033
ARMA	535,833.8	462,131	73,702.82	95	7,75.8192
CSTS	413,830.8	618,588.4	−204,758	95	−2,155.34
MEAN	454,006.3	1,714,191	−1,260,185	95	−13,265.1
(b) 2011					
MAV5	840,926.2	706,118.6	134,807.6	152	886.8921
EMA5	−794,276	784,070.1	−1,578,346	152	−10,383.9
LIV	2,433,862	**500,012.3**	1,933,850	152	12,722.7
RGN	**3,170,605**	1,090,987	**2,079,618**	152	**13,681.7**
ARMA	679,499.3	786,602.5	−107,103	152	−704.63
CSTS	−4,119,967	600,665.3	−4,720,633	152	−31,056.8
MEAN	4,168,410	2,752,602	1,415,807	152	9,314.522

The holdout period is the last 20 days of each S&P 500 index futures option contracts. There are seven IV estimates for the trading rule test: MAV5 is the 5-day moving averages method, EMA5 is the 5-day exponential moving averages method, LIV is Previous IV method, RGN is the Regression method, ARMA is autoregressive-moving-average model, CSTS is the cross-sectional time-series predictive regression model represented in Eq. (16.16) where using ARMA as predictor method, and MEAN is the constant value over the entire period equal to the mean of individual IV series. The definitions of first five IV estimates are indicated in Tables 16.2 and 16.3. The gross value of all trades are included the bought and sold price of options plus the value in the end of maturity if the trades are not closed out before maturity. Total trading costs are included the total transaction fees, margin interest costs, future liquidity costs, and option liquidity costs. The net value of all equals to gross value of all trades minus total trading costs. The net profit or loss per trade is the value of net value of all trades divided by number of trade

utilized some of the insights of time-series analysis as would be impounded in the optimal time-series predictors, ARMA model. Also, it takes into account the historical 5-days and 20-days standard deviation of the continuous return for the underlying futures contract, the short-term variability and skewness and kurtosis of the IV, the time-to-maturity, and weekday effects. Table 16.3a, b summarize the cumulative trading results for the selected options contract in Table 16.1. For both years, EMA5, LIV, and RGN perform better than the sophisticated model such as cross-sectional time-series predictive regression. The results implied that the ARMA model may have over-fitting problem and thus make CSTS model perform worse. The worse prediction is using MEAN model to estimate IV. MEAN model's IV is constant for entire period of contract; thus, MEAN model neither deal with the fluctuation of option market nor response to everyday's new important information. It also implied that the constant volatility setting in BSM model may be misspecified.

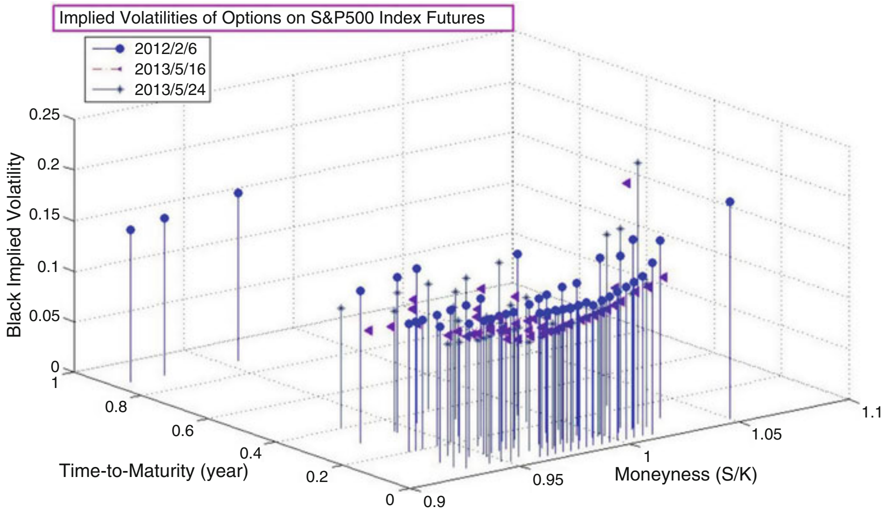


Fig. 16.1 Implied volatilities in black model

16.4.4 Structural Parameter Estimation and Performance of CEV Model

In Fig. 16.1, we find that each contract's Black IV varies across moneyness and time-to-maturity. This graph shows volatility skew (or smile) in options on S&P 500 index futures, i.e., the implied volatilities decrease as the strike price increases (the moneyness level decreases).

Even though everyday implied volatility surface changes, this characteristic still exists. Therefore, we divided future option contracts into a six by four matrix based on moneyness and time-to-maturity levels when we estimate implied volatilities of futures options in CEV model framework in accordance with this character. The whole option samples expired within the period of 2010 to 2013 contains 30,364 observations. The whole period of option prices is from March 19, 2009 to November 5, 2013. The observations for each group are presented in Table 16.4.

Since most trades are in the futures options with short time-to-maturity, the estimated implied volatility of the option samples in 2009 may be significantly biased because we did not collect the futures options expired in 2009. Therefore, we only use option prices in the period between January 1, 2010 and November 5, 2013 to estimate parameters of CEV model. In order to find global optimization instead of local minimum of absolute pricing errors, the ranges for searching suitable δ_0 and α_0 are set as $\delta_0 \in [0.01, 0.81]$ with interval 0.05, and $\alpha_0 \in [-0.81,\ 1.39]$ with interval 0.1, respectively. First, we find the value of parameters, $(\widehat{\delta_0}, \widehat{\alpha}_0)$, within the ranges such that minimize value of absolute pricing errors in Eq. (16.19). Then we use this pair of parameters, $(\widehat{\delta_0}, \widehat{\alpha}_0)$, as optimal initial estimates in the procedure of

Table 16.4 Average daily and total number of observations in each group

Time-to-maturity(TM)	$TM < 30$		$30 \leqq TM \leqq 60$		$TM > 60$		All TM	
Moneyness (S/K ratio)	Daily Obs.	Total Obs.	Daily Obs.	Total Obs.	Daily Obs.	Total Obs.	Daily Obs.	Total Obs.
S/K ratio >1.01	1.91	844	1.64	499	1.53	462	2.61	1, 805
$0.98 \leqq$ S/K ratio $\leqq 1.01$	4.26	3, 217	2.58	1, 963	2.04	1, 282	6.53	6, 462
$0.95 \leqq$ S/K ratio < 0.98	5.37	4, 031	3.97	3, 440	2.58	1, 957	9.32	9, 428
$0.9 \leqq$ S/K ratio < 0.95	4.26	3, 194	4.37	3, 825	3.27	2, 843	9.71	9, 862
S/K ratio <0.9	2.84	764	2.68	798	2.37	1, 244	4.42	2, 806
All Ratio	12.59	12, 050	10.78	10, 526	7.45	7, 788	27.62	30, 364

The whole period of option prices is from March 19, 2009 to November 5, 2013. Total observations is 30, 364. The lengths of period in groups are various. The range of lengths is from 260 (group with ratio below 0.90 and time-to-maturity within 30 days) to 1,100 (whole samples)

Table 16.5 Initial parameters of CEV model for estimation procedure

Time-to-Maturity (TM)	TM < 30		30 ≦ TM ≦ 60		TM > 60		All TM	
Moneyness (S/K ratio)	α_0	δ_0.	α_0	δ_0	α_0	δ_0	α_0	δ_0
S/K ratio >1.01	0.677	0.400	0.690	0.433	0.814	0.448	0.692	0.429
0.98 ≦ S/K ratio ≦ 1.01	0.602	0.333	0.659	0.373	0.567	0.361	0.647	0.345
0.95 ≦ S/K ratio < 0.98	0.513	0.331	0.555	0.321	0.545	0.349	0.586	0.343
0.9 ≦ S/K ratio < 0.95	0.502	0.344	0.538	0.332	0.547	0.318	0.578	0.321
S/K ratio <0.9	0.777	0.457	0.526	0.468	0.726	0.423	0.709	0.423
All Ratio	0.854	0.517	0.846	0.512	0.847	0.534	0.835	0.504

The sample period of option prices is from January 1, 2010 to November 5, 2013. During the estimating procedure for initial parameters of CEV model, the volatility for S&P 500 index futures equals to $\delta_0 S^{\alpha_0-1}$

Table 16.6 Total number of observations and trading days in each group

Time-to-Maturity (TM)	TM < 30		30 ≦ TM ≦ 60		TM > 60		All TM	
Moneyness (S/K ratio)	Days	Total Obs.	Days	Total Obs.	Days	Total Obs.	Days	Total Obs.
S/K ratio >1.01	172	272	104	163	81	122	249	557
0.98≦S/K ratio≦1.01	377	1,695	354	984	268	592	448	3,271
0.95≦S/K ratio < 0.98	362	1,958	405	1,828	349	1,074	457	4,860
0.9≦S/K ratio < 0.95	315	919	380	1,399	375	1,318	440	3,636
S/K ratio <0.9	32	35	40	73	105	173	134	281
All Ratio	441	4,879	440	4,447	418	3,279	461	12,605

The subsample period of option prices is from January 1, 2012 to November 5, 2013. Total observations is 13, 434. The lengths of period in groups are various. The range of lengths is from 47 (group with ratio below 0.90 and time-to-maturity within 30 days) to 1,100 (whole samples). The range of daily observations is from 1 to 30

estimating local minimum minAPE based on steps (1)–(3) in Sect. 16.3.2.3. To compare with the option pricing performance of Black model, we set the interval between 0.01 and 0.08 to find optimal implied volatility via estimation procedure in Sect. 16.3.2.1. The initial parameter setting of CEV model is presented in Table 16.5.

In Table 16.5, the average sigma are almost the same while the average alpha value in either each group or whole sample is less than one. This evidence implies that the alpha of CEV model can capture the negative relationship between S&P 500 index future prices and its volatilities shown in Fig. 16.1. The instant volatility of S&P 500 index future prices equals to $\delta_0 S^{\alpha_0-1}$ where S is S&P 500 index future prices, δ_0 and α_0 are the parameters in CEV model. The estimated parameters in Table 16.9 are similar across time-to-maturity level but volatile across moneyness.

Because of the implementation and computational costs, we select the sub-period from January 2012 to November 2013 to analyze the performance of CEV model. The total number of observations and the length of trading days in each group are presented in Table 16.6. The estimated parameters in Table 16.7 are similar across time-to-maturity level but volatile across moneyness. Therefore, we investigate the

Table 16.7 Average daily parameters of in-sample

Time-to-Maturity (TM)	TM < 30				30≦TM≦60				TM > 60				All TM			
Moneyness (S/K ratio)	CEV			Black	CEV			Black	CEV			Black	CEV			Black
Parameters	α	δ	IV	IV	α	δ	IV	IV	α	δ	IV	IV	α	δ	IV	IV
S/K ratio >1.01	0.29	0.19	0.188	0.200	0.14	0.18	0.183	0.181	0.29	0.21	0.204	0.196	0.25	0.19	0.1890	0.1882
0.98≦S/K ratio≦1.01	0.34	0.16	0.162	0.1556	0.30	0.16	0.154	0.147	0.14	0.16	0.155	0.155	0.39	0.17	0.151	0.150
0.95≦S/K ratio < 0.98	0.22	0.13	0.137	0.135	0.30	0.13	0.134	0.131	0.24	0.14	0.141	0.139	0.37	0.14	0.136	0.132
0.9≦S/K ratio < 0.95	0.05	0.15	0.159	0.152	0.25	0.13	0.133	0.128	0.26	0.14	0.136	0.131	0.38	0.14	0.135	0.129
S/K ratio <0.9	−0.23	0.22	0.252	0.243	−1.67	0.14	0.193	0.159	0.25	0.15	0.145	0.142	0.23	0.15	0.157	0.152

The in-sample period of option prices is from January 1, 2012 to May 30, 2013. In the in-sample estimating procedure, CEV implied volatility for S&P 500 index futures (CEV IV) equals to $\delta(S/K \ \text{ratio})^{\alpha-1}$ in according to reduce computational costs. The optimization setting of finding CEV IV and Black IV is under the same criteria performance of all groups except the groups on the bottom row of Table 16.8. The performance of models can be measured by either the implied volatility graph or the average absolute pricing errors (AveAPE). The implied volatility graph should be flat across different moneyness level and time-to-maturity. We use subsample like Bakshi et al. (1997) and Chen et al. (2009) did to test implied volatility consistency among moneyness-maturity categories. Using the subsample data from January 2012 to May 2013 to test in-the-sample fitness, the average daily implied volatility of both CEV and Black models, and average alpha of CEV model are computed in Table 16.7. The fitness performance is shown in Table 16.8. The implied volatility graphs for both models are shown in Fig. 16.2. In Table 16.7, we estimate the optimal parameters of CEV model by using a more efficient program. In this efficient program, we scale the strike price and future price to speed up the program where the implied volatility of CEV model equals to $\delta\left(\text{ratio}^{\alpha-1}\right)$, ratio is the moneyness level, δ and α are the optimal parameters of program which are not the parameters of CEV model in Eq. (16.17). In Table 16.8, we found that CEV model perform well at in-the-money group.

Figure 16.2 shows the IV computed by CEV and Black models. Although their implied volatility graphs are similar in each group, the reasons to cause volatility smile are totally different. In Black model, the constant volatility setting is misspecified. The volatility parameter of Black model in Fig. 16.2b varies across moneyless and time-to-maturity levels while the IV in CEV model is a function of the underlying price and the elasticity of variance (alpha parameter). Therefore, we can image that the prediction power of CEV model will be better than Black model because of the explicit function of IV in CEV model. We can use alpha to measure the sensitivity of relationship between option price and its underlying asset. For example, in Fig. 16.2c, the in-the-money future options near expired date have significantly negative relationship between future price and its volatility.

The better performance of CEV model may result from the over-fitting issue that will hurt the forecastability of CEV model. Therefore, we use out-of-sample data from June 2013 to November 2013 to compare the prediction power of Black and CEV models. We use the estimated parameters in previous day as the current day's input variables of model. Then, the theoretical option price computed by either Black or CEV model can calculate bias between theoretical price and market price. Thus, we can calculate the average absolute pricing errors (AveAPE) for both models. The lower value of a model's AveAPE, the higher pricing prediction power of the model. The pricing errors of out-of-sample data are presented in Table 16.9. Here we find that CEV model can predict options on S&P 500 index futures more precisely than Black model. Based on the better performance in both in-sample and out-of-sample, we claim that CEV model can describe the options of S&P 500 index futures more precisely than Black model.

Table 16.8 AveAPE performance for in-sample-fitness

Time-to-Maturity (TM)	TM < 30			30≦TM≦60			TM > 60			All TM		
Moneyness (S/K ratio)	CEV	Black	Obs.	CEV	Black	Obs.	CEV	Black	Obs.	CEV	Black	Obs.
S/K ratio >1.01	**1.65**	1.88	202	1.81	**1.77**	142	5.10	**5.08**	115	5.80	**6.51**	459
0.98 ≦ S/K ratio ≦ 1.01	**6.63**	7.02	1290	**4.00**	4.28	801	4.59	**4.53**	529	**18.54**	18.90	2620
0.95 ≦ S/K ratio < 0.98	2.38	**2.34**	1560	4.25	**4.14**	1,469	3.96	**3.89**	913	14.25	**14.15**	3942
0.9 ≦ S/K ratio < 0.95	0.69	**0.68**	710	1.44	**1.43**	1,094	3.68	**3.62**	1,131	**7.08**	7.10	2935
S/K ratio <0.9	0.01	0.01	33	**0.13**	0.18	72	0.61	**0.60**	171	0.69	**0.68**	276

The in-sample period of option prices is from January 1, 2012 to May 30, 2013

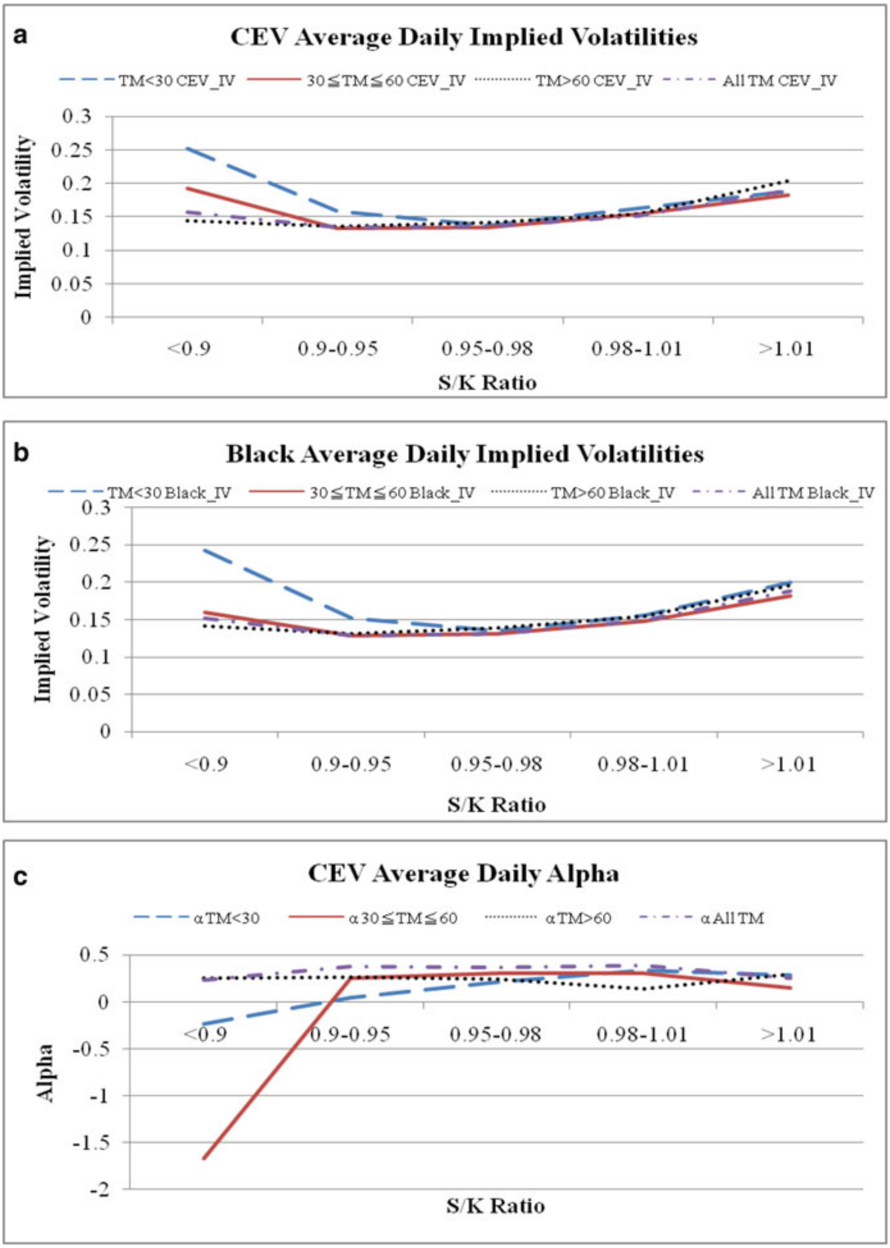


Fig. 16.2 Implied volatilities and CEV Alpha Graph

Table 16.9 AveAPE performance for out-of-sample

Time-to-Maturity (TM)	TM < 30		30 ≦ TM ≦ 60		TM > 60		All TM	
Moneyness (S/K ratio)	CEV	Black	CEV	Black	CEV	Black	CEV	Black
S/K ratio >1.01	**3.22**	3.62	**3.38**	4.94	**8.96**	13.86	**4.25**	5.47
0.98 ≦ S/K ratio ≦ 1.01	**2.21**	2.35	2.63	**2.53**	**3.47**	3.56	**2.72**	2.75
0.95 ≦ S/K ratio < 0.98	**0.88**	1.04	**1.42**	1.46	1.97	**1.95**	**1.44**	1.45
0.9 ≦ S/K ratio < 0.95	**0.34**	0.53	**0.61**	0.62	1.40	1.40	**0.88**	0.90
S/K ratio <0.9	**0.23**	0.79	**0.25**	0.30	1.28	**1.27**	**1.03**	1.66

16.5 Conclusion

The purpose of this essay has been to improve the interpretation and forecasting of individual implied volatility (IV) for call options on S&P500 index futures in 2010 to 2013. The two alternative methods used in this essay are cross-sectional time-series analysis and CEV model. These two alternative approaches give different perspective of estimating IV. The cross-sectional time-series analysis focuses on the dynamic behavior of volatility in each option contracts and captures other trading behaviors such as week effect and in/out of the money effect. On the other hand, CEV model generalizes implied volatility surface as a function of asset price.

By empirically explaining the composition through time-series analysis and cross-sectional time-series regression models, the disadvantages to evaluating an option IV by Black model have been demonstrated. More importantly, the results based on our trading strategy provide some evidence as to how the Black option pricing model might be misspecified, or jointly, how the market might be inefficient. Though the original model implicitly assumes a frictionless market and a constant volatility term, market realities along with past studies would not be able to substantiate these types of assumptions. The forecasting performances of seven time-series regression models based on our trading strategy show that the simple regression models perform better than sophisticated cross-sectional time-series models because of over-fitting problem in the advanced models. In addition, although our trading rules based on the prediction of these models can make profit, the net profit depends on the transaction costs. Therefore, the setting of trading strategy should be necessarily adjusted to the transaction costs.

We also show that CEV model performs better than Black model in aspects of either in-sample fitness or out-of-sample prediction. The setting of CEV model is more reasonable to depict the negative relationship between S&P 500 index future price and its volatilities. The elasticity of variance parameter in CEV model captures the level of this characteristic. The stable volatility parameter in CEV model in our empirical results implies that the instantaneous volatility of index future is mainly determined by current future price and the level of elasticity of variance parameter.

In sum, we suggest predict individual option contract by using simple regression analysis instead of advanced cross-sectional time-series model. Even though the moneyness and week effect have significant influence on index future option prices, the over-fitting problem in an advanced cross-sectional time-series model will decrease its pricing forecastability. With regard to generate implied volatility surface to capture whole prediction of the future option market, the CEV model is the better choice than Black model because it not only captures the skewness and kurtosis effects of options on index futures but also has less computational costs than other jump-diffusion stochastic volatility models. In future research, we can apply CEV model and its Greek measures to other liquid option markets to test market efficiency based on our trading rules.

References

Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics, 21*, 243–247.

Akaike, H. (1977). On entropy maximization principle. In *Applications of statistics*. P.R. Krishnaiah (Ed.), (pp. 27–41) Amsterdam, North Holland.

Ansley, C. F., & Newbold, P. (1980). Finite sample properties of estimators for autoregressive moving average models. *Journal of Econometrics, 13*, 159–183.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*, 2003–2049.

Beckers, S. (1980). The constant elasticity of variance model and its implications for option pricing. *Journal of Finance, 35*, 661–673.

Benton, D., & Krishnamoorthy, K. (2003). Compute discrete mixtures of continuous distributions: Noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient. *Computational Statistics and Data Analysis, 43*, 249–267.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal, 31*, 36–41. 61–72.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–659.

Brenner, M., Courtadon, G., & Subrahmanyam, M. (1985). Options on the spot and options on futures. *Journal of Finance, 40*, 1303–1317.

Brenner, M., & Galai, D. (1981). The properties of the estimated risk of common stocks implied by option prices. *Working paper*, University of California at Berkeley.

Chance, D. M. (1986). Empirical tests of the pricing of index call options. *Advances in Futures and Options Research, 1*, 141–166.

Chen, R., & Lee, C. F. (1993). A constant elasticity of variance family of stock prices in option pricing: Review and integration. *Journal of Financial Studies, 1*, 29–51.

Chen, R., Lee, C. F., & Lee, H. (2009). Empirical performance of the constant elasticity variance option pricing model. *Review of Pacific Basin Financial Markets and Policies, 12*(2), 177–217.

Chiras, D., & Manaster, S. (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics, 6*, 213–234.

Cox, J. C. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. *Working paper*, Stanford University.

Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics, 3*, 145–166.

Day, T. E., & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics, 52*, 267–287.

DelBaen, F., & Sirakawa, H. (2002). A note on option pricing for the constant elasticity of variance model. *Asia-Pacific Financial Markets, 9*, 85–99.

Ding, C. G. (1992). Algorithm AS275: Computing the non-central χ^2 distribution function. *Applied Statistics, 41*, 478–482.
Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *Journal of Finance, 53*, 2059–2106.
Emanuel, D. C., & MacBeth, J. D. (1982). Further results on the constant elasticity of variance call option pricing model. *Journal of Financial and Quantitative Analysis, 17*, 533–554.
Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica, 50*, 987–1008.
Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance, 5*, 317–345.
Fleming, J., Ostdiek, B., & Whaley, R. E. (1996). Trading costs and the relative rates of price discovery in stock, futures, and option markets. *Journal of Futures Markets, 16*, 353–387.
Geske, R. (1979). The valuation of compound options. *Journal of Financial Economics, 7*, 63–81.
Hannan, E. J., & Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika, 69*, 81–94.
Harvey, C. R., & Whaley, R. E. (1991). S&P 100 index option volatility. *Journal of Finance, 46*, 1551–1561.
Harvey, C. R., & Whaley, R. E. (1992a). Market volatility prediction and the efficiency of the S&P 100 index option market. *Journal of Financial Economics, 31*, 43–73.
Harvey, C. R., & Whaley, R. E. (1992b). Dividends and S&P 100 index option valuation. *Journal of Futures Market, 12*, 123–137.
Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–343.
Hillmer, S. C., & Tiao, G. C. (1979). Likelihood function of stationary multiple autoregressive moving average models. *Journal of the American Statistical Association, 74*, 652–660.
Hsu, Y. L., Lin, T. I., & Lee, C. F. (2008). Constant elasticity of variance (CEV) option pricing model: Integration and detailed derivation. *Mathematics and Computers in Simulation, 79*, 60–71.
Hull, J. C. (2011). *Options, futures, and other derivatives* (8th ed.). Englewood Cliffs, NJ: Prentice-Hall.
Hull, J. C., & White, A. (1987). The pricing of options on assets with stochastic volatility. *Journal of Finance, 42*, 281–300.
Jackwerth, J. C., & Rubinstein, M. (2001). Recovering stochastic processes from option prices. Working paper, London Business School.
Jarrow, R., & Rudd, A. (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics, 10*, 347–369.
Jarrow, R., & Rudd, A. (1983). Tests of an approximate option valuation formula. In M. Brenner (Ed.), *Option pricing: Theory and applications* (pp. 81–100). Lexington, MA: Lexington Books.
Jones, C. S. (2003). The dynamics of stochastic volatility: Evidence from underlying and option markets. *Journal of Econometrics, 116*, 181–224.
Larguinho, M., Dias, J. C., & Braumann, C. A. (2013). On the computation of option prices and Greeks under the CEV model. *Quantitative Finance, 13*(6), 907–917.
Latane, H. A., & Rendleman, R. J. (1976). Standard deviations of stock price ratios implied in option prices. *Journal of Finance, 31*, 369–381.
Lee, R. W. (2001). Implied and local volatilities under stochastic volatility. *International Journal of Theoretical and Applied Finance, 4*, 45–89.
Lee, C. F., & Wu, C. (1985). The impacts of kurtosis on risk stationarity: Some empirical evidence. *Financial Review, 20*, 263–269.
Lee, C. F., Wu, T., & Chen, R. (2004). The constant elasticity of variance models: New evidence from S&P 500 index options. *Review of Pacific Basin Financial Markets and Policies, 7*(2), 173–190.
Lewis, A. S. (2000). *Option Valuation under stochastic volatility with Mathematica code*. NewportBeach, CA: Finance Press.

Macbeth, J. D., & Merville, L. J. (1979). An empirical examination of the Black-Scholes call option pricing model. *Journal of Finance, 34*, 1173–1186.

MacBeth, J. D., & Merville, L. J. (1980). Tests of the Black-Scholes and Cox Call option valuation models. *Journal of Finance, 35*, 285–301.

Medvedev, A., & Scaillet, O. (2007). Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility. *Review of Financial Studies, 20*, 427–459.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management science, 4*, 141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3*, 125–144.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics, 45*, 267–290.

Park, H. Y., & Sears, R. S. (1985). Estimating stock index futures volatility through the prices of their options. *Journal of Futures Markets, 5*, 223–237.

Phillips, S. M., & Smith, C. W. (1980). Trading costs for listed options: The implications for market efficiency. *Journal of Financial Economics, 8*, 179–201.

Pun, C. S., & Wong, H. Y. (2013). CEV asymptotics of American options. *Journal of Mathematical Analysis and Applications, 403*(2), 451–463.

Ramaswamy, K., & Sundaresan, S. M. (1985). The valuation of options on futures contracts. *Journal of Finance, 40*, 1319–1340.

Rubenstein, M. (1985). Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978. *Journal of Finance, 40*, 455–480.

Schmalensee, R., & Trippi, R. (1978). Common stock volatility estimates implied by option premia. *Journal of Finance, 33*, 129–147.

Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *Journal of Finance, 44*(1), 211–219.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative Analysis, 22*, 419–438.

Scott, L. O. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rate: Application of Fourier inversion methods. *Mathematical Finance, 7*(4), 413–424.

Singh, V. K., & Ahmad, N. (2011). Forecasting performance of constant elasticity of variance model: Empirical evidence from India. *International Journal of Applied Economics and Finance, 5*, 87–96.

Stein, E. M., & Stein, J. C. (1991). Stock price distribution with stochastic volatility: An analytical approach. *Review of Financial Studies, 4*, 727–752.

Stephan, J. A., & Whaley, R. E. (1990). Intraday price change and trading volume relations in the stock and stock option markets. *Journal of Finance, 45*, 191–220.

Tsai, W. C. (2014). Improved method for static replication under the CEV model. *Finance Research Letters, 11*, 194–202.

Whaley, R. E. (1981). On the valuation of American call options on stocks with known dividends. *Journal of Financial Economics, 9*, 207–211.

Whaley, R. E. (1982). Valuation of American call options on dividend-paying stocks: Empirical tests. *Journal of Financial Economics, 10*, 29–58.

Wiggins, J. B. (1987). Option values under stochastic volatilities. *Journal of Financial Economics, 19*, 351–372.

Wolf, A. (1982). Fundamentals of commodity options on futures. *Journal of Futures Markets, 2*, 391–408.

Chapter 17
The Swiss Black Swan Unpegging Bad Scenario: The Losers and the Winners

Sebastien Lleo and William T. Ziemba

JEL codes: B 41, C12, C52, G01, G11

17.1 The Swiss Franc Peg

The Swiss National Bank (SNB) pegged the Swiss franc (CHF) to the euro at 1.20 on September 6, 2011, thereby tracking the euro in its moves against all other currencies. The peg was adopted in the midst of the European debt crisis as the Swiss currency experienced massive safe haven inflows. These flows of funds were both threatening the competitiveness of the Swiss economy and creating significant asset bubbles within Switzerland, notably property. In pegging, the SNB moved

Without implicating them, thanks are due to Alex Ziegler, Thorsten Hens and the cited sources for help in the analysis in this paper. We are particularly indebted to Rachel Ziemba and Sandra Schwartz for their timely advice, careful analysis, and thoughtful suggestions. Forthcoming in J. Guerard, Ed (2015) *Portfolio construction, measurement and efficiency: Essays in honor of Jack Treynor*, Springer. This is a substantially revised version of Lleo and Ziemba (2015).

S. Lleo (✉)
Finance Department, NEOMA Business School, Reims, France
e-mail: sebastien.lleo@NEOMA-bs.fr

W.T. Ziemba
Financial Modeling and Stochastic Optimization (Emeritus), University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Systemic Risk Centre, London School of Economics, Houghton Street, London WC2A 2AE, UK
e-mail: wtzimi@mac.com

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_17

away from merely intervening in the currency markets to a defined target franc-euro exchange rate. However, both approaches would have significantly expanded their balance sheet.

This was an agreement to buy euros at that rate even with a growing desire in the market to sell euros which, unpegged, would have driven the euro-franc rate down. This meant that the SNB was pledging to printing Swiss francs on demand, to buy large amounts of other currencies. According to Table 17.1 the SNB "lost" about CHF78 billion, which is about 12 % of the Swiss GDP. They made some CHF38 billion during 2014.

The dual purposes of the SNB are to make money for its investors (about 55 % of the SNB's shares are held by public institutions such as cantons, while the remaining 45 % are openly traded on the stock market) and to use its balance sheet as a means of meeting its monetary policy goals including fighting deflation; see Article 5 of the Federal Act on the Swiss National Bank (The Federal Assembly of the Swiss Confederation, March 2012). This unique structure of the SNB, compared to most other Central Banks, meant that the two mandates conflicted preventing the SNB from achieving its monetary policy goals.

Central bank reserve accumulation tends to be negative for its returns, since they have to effectively take on a negative carry, or accept the risk of future losses. If their accumulation was the main reason why the CHF was staying weak at some point, they would have to take losses.

Moreover, the recent gold referendum (even though it failed) was a sign that the political costs of expanding the balance sheet had increased and that the Swiss public might not have been happy about an increase in the fiscal cost of keeping the floor. A sharper move of the European Central bank (ECB) toward Quantitative Easing (QE), a return of the Eurozone (EZ) crisis in the name of Greece or Russia related uncertainties, effectively could put pressure on the exchange rate and thus require it to expand its balance sheet even more by buying a lot of euros.

The initial aim of the EUR/CHF floor/peg was about to even more appreciation of the CHF which was being buffeted by capital inflows during the euro crisis and more recently. However, the conflict between the SNB's mandates and the fact that the Swiss were questioning the effectiveness of the peg made the political costs of maintaining the peg too high to bear.

Arguably, Switzerland is not the only country directly exposed to the variation of the Euro and to fundamental changes in the ECB's policies. Denmark, whose Krone is pegged to the Euro,[1] was forced to cut interest four times in just 18 days to defend the fixed exchange rate. Their move, on February 5th, 2015, made the headline by slashing the Danmarks Nationalbank's interest rate on certificates of deposit by 0.25 % points to a negative interest rate of −0.75 %. In an effort to defend

---

[1]The Danish Krone was part of the original European Exchange Rate Mechanism (ERM) in force before the creation of the Euro. Following a referendum in 2000, which saw a rejection of the Euro, Denmark kept its Krone but pegged it closely to the Euro within an updated version of the ERM, called ERM II. Denmark is currently the only country left in ERM II after Greece officially adopted the Euro in 2001. While the ERM II officially allows currencies to float within a range of +/−15 % with respect to the euro, Denmark has opted for a narrow +/−2.25 % band.

Table 17.1 Approximate loss of the SNB in the two days following the lifting of the peg

Currency	US$	EUR	JPY	GBP	CAD	Other	Total
Amount Sept 30 (in mln)	142,366	174,335	4,490,747	21,500	24,492	31,578	
CHF Equivalent Sept 30 (in mln)	136,102	210,335	39,164	33,336	20,938	31,578	471,453
% of Currency Reserves	28.87 %	44.61 %	8.31 %	7.07 %	4.44 %	6.70 %	100.00 %
FX Rate Jan 14	1.0187	1.20095	0.008682	1.5519	0.8524	1	
FX Rate Jan 16	0.8587	0.9941	0.0073	1.3016	0.7168	0.85	
FX Return	−15.71 %	−17.22 %	−15.92 %	−16.13 %	−15.91 %	−15.00 %	
MTM Gain/Loss (in CHF mln)	−22,779	−36,061	−6,206	−5,381	−3,321	−4,737	−78,485

Source: Private correspondence with Alex Ziegler, Jan 17, 2015, calculations based on late September published holdings the peg, the central bank's currency reserves have increased close to 100 % in recent months. They now amount to more than USD 110 billions, which is about one third of Denmark's 2014 GDP.[2]

Swiss policy is a form of QE but because Swiss is a small open economy, rather than printing money directly, they committed to buy Euros to weaken the CHF.

These costs were deemed too great, and the SNB is now relying on negative interest rates (−0.75 %) to try to weaken CHF vs EUR. It is not clear if this will be effective. Moreover, there could be a lot of collateral damage if investors avoid putting on any hedges since they trusted the floor and policy.

The SNB did a poor job of managing expectations and the risk is that they will end up having made a policy mistake since they are even more mired in deflation, but there were political pressures.

If the SNB was just focused on reflationary policies they would have implemented policies, that yield a weaker exchange rate. What surprised market actors was the fact that if anything with US$/EUR moving so much in anticipation of ECB QE, and the weaker oil price adding to global/european deflationary trends, the EUR/CHF should have weakened.

In early February the SNB unofficially began targeting an exchange rate *corridor* of 1.05–1.10 CHF/euro. The bank was willing to incur losses of a further CHF10 billion over a period of time that they did not specify. This was another non-transparent action by the SNB. But it amounts to a revaluation of the Swiss franc by 10–15 %.

17.2 Why Did the SNB Start the Peg and Why Did They Eliminate It?

Why They Did It The peg was installed on September 6, 2011 (Swiss National Bank 2011). The trigger for the SNB's decision to end the peg so precipitously, following an earlier announcement in the same week that they were maintaining it, was probably the ECB's hints that it was ready to announce a large scale program of quantitative easing to attempt to move the EuroZone out of deflation. Indeed, the ECB officially announced the details of its QE programme on January 22, 2015, one week exactly after Switzerland removed its peg. The programme calls for a €60 billion monthly bond purchase, totalling about €1 trillion to start in March 2015 until September 2016.

Could the SNB have eliminated the peg gradually? We don't think so. They could have announced it over a weekend to soften the blow but the final effect would have been similar.

The peg was effectively fixed at 1.20. Hong Kong is in a similar situation as its currency is pegged to the US$.

[2] We thank Rachel Ziemba for providing us with this data.

During the Asian financial crisis, the Hong Kong Monetary Authority (HKMA) was among the first central banks to engage in quantitative easing. The HKMA was willing to expand its balance sheet massively, including by purchasing local equity.

Both Hong Kong and Switzerland have small economies with the peg against a much larger economy. Both have been affected by the larger economy's monetary policy. Both are financial centres. This situation can result in shocks. In the case of Hong Kong, the attempt to balance between the USA and China will complicate the maintenance of the peg over the longer term. For Switzerland, the shock was the prospect of the ECB's QE. Krugman (2015a) argues that the situations of Switzerland and Hong Kong are different: "the institutional setup and history of Hong Kong plays every differently with the hard-money ideology than the Swiss peg did . . . ." Hong Kong has a currency board to maintain the peg, and the HKMA does not have the mixed ownership structure of the SNB. The Swiss could have maintained the peg forever but it was nagging from hard money types that led to the change in priority. The Swiss currency intervention was the result of a huge expansion of the central bank's balance sheet and printing money even if the goal was to keep it from getting stronger.

17.3 How Does Quantitative Easing Work and What Are It Is Costs and Benefits?

To see why talks about an ECB-led Quantitative Easing programme most probably delivered the final blow to the Swiss peg, we need to understand how quantitative easing work and what its effects have been so far.

Although Quantitative Easing is new in the Eurozone, it has been used for more than 6 years in the US. The track record of QE in the USA is checkered at best. QE managed to relfate asset markets while failing to support final demand and pushing investors into higher yielding assets. According to Sandra Schwartz (2015):

> We have been sold a bill of goods on quantitative easing. It is not a new monetary policy, but a variant, only the Instruments are different - buying up the debt of banks and others rather than buying up government debt.
>
> With quantitative easing all economic policy has been placed on the shoulders of monetary policy and this has been very indirect and has led primarily to an asset bubble.
>
> With fiscal policy, that is running a government deficit, good and services are bought directly. Quantitative easing happens when the bank buys the debt that the government has created to facilitate it. When there is unemployment this does not crowd out private investment but gets the economy moving. It could build infrastructure that directly will later help private investment. As it is directly spent on goods and services, it does not create an asset bubble but it creates jobs, employment and income.

The USA took an approach to stimulate the economy and not have austerity measures imposed on the companies and people. This has led to a robust economy and a stock market up three times since the March 6 low. Figure 17.1 displays the evolution of the S&P500 since the start of QE 1 in late November 2008. The level of

Fig. 17.1 Effect of quantitative easing on the US equity market: evolution of the S&P 500 between November 15, 2008 and September 21, 2015

the S&P500 has increased by 144 % between November 15th, 2008 and its peak on May 21st, 2015. At the time of writing, the Federal Reserve has stopped purchasing assets and the S&P 500 has declined, is still 122 % above its November 15th, 2008 level. The Federal Reserve now faces the decision of how to deal with a balance sheet worth US$ 4 trillion balance sheet: should it start unwinding its positions or simply wait until the bonds it has purchased mature? Neither exit strategies are expected to cause much trouble on the financial markets.

Europe faces a complex situation with more players. Europe took a different route—austerity. This has caused trouble in many places, pushing unemployment among the youths to between 25 % and 50 %. The worst case is Greece and, in 2015, there was a lot of trouble. The 60 billion per month, out of some 1.2 trillion bond buying, will have to delicately balance inflation and deflation. A serious negotiation is taking place between Germany, who as a major exporter benefits from a situation where the EuroZone does not implode, and Greece, who cannot continue with the heavy austerity. While both countries have vastly different political and economic structures, they both benefit from a weaker Euro. So a political compromise ironed out but the states and a quantitative easing sponsored by the European Central Bank are coming at the same time to the Eurozone. With the Euro weakening, the anticipation of a sharp increase of foreign capitals flowing into Switzerland triggered the Swiss currency move.

The real problem with quantitative easing in the USA and Europe is where the money goes: to banks, whereas channeling it to people would be more effective. The expectation has been that bank would reorientate toward their historical financial intermediation role after the credit crisis of 2007–2009. Chuptka (2015) following Peter Schiff of Euro Capital argues as we to do that unemployment is the key problem.

17.4 The Currency Moves

Currencies tend to trend and reverse sharply. A typical example is the US$/EUR from 2002 to 2007 when the euro gradually fell until it sharply reversed. The trade to sell puts out of the money on the US$/EUR exchange rate was very successful but it ended badly when the currency turned. For example, puts that were 2 cents one day, 4 cents the next day, and then 28 cents the next ended up at $4. Figure 17.2 shows the Euro exchange rate from its start on January 1, 1999 to September 21, 2015. Physical Euros coins and banknotes came in use across the Eurozone on January 1, 2002. Currently, the Eurozone includes the following 19 of the 28 member states of the European Union: Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia, and Spain.

Usually currencies move 0.5–1 % or even 2 % per day on a large move. The SNB's drop of the peg caused an immediate 39 % increase in the CHF vs the EUR. See the January 15, 2015 move in Fig. 17.3.

Figure 17.4 displays the evolution of the Swiss Franc against the four major currencies: Euro (EUR), British Pound (GBP), US Dollar (US$), and Japanese Yen (JPY).

Figure 17.5 shows the evolution of the Swiss Franc against the other European currencies: Danish Krone (DKK), Norwegian Krone (NOK), Czech Koruna (CZK), Hungarian Forint (HUF), Polish Zloty (PLN), Russian Ruble (RUB), Swedish Krona (SEK).

Figure 17.6 shows the evolution of the Swiss Franc against the currencies of commodity producing countries: Canadian Dollar (CAD), Brazilian Real (BRL), South African Rand (ZAR), Australian Dollar (AUD), and New Zealand Dollar (NZD).

Figure 17.7 displays the relative performance of the CHF/EUR exchange rate against the Swiss equity market index, the Swiss Market Index (SMI), and the SNB's stock price from January 1, 2014 to September 20, 2015. In the aftermath of the January 15th decision to remove the peg, the SNB's stock price fared far better than the swiss market and the exchange rate. The Swiss stock market and SNB stock price have recovered: they are both up 6.5 % over their December 1st level. On the other hand, the Swiss franc is still 11.3 % lower than on December 1st against the euro.

Fig. 17.2 The US$/EUR exchange rate from January 1, 1999 to September 21, 2015. The Euro became an effective currency on January 1, 2002 when the first coins and banknotes started circulating (closing EUR/USD rate on December 31, 2001= 0.9038, source: ECB)

Figure 17.8 shows the performance of the SMI versus the DAX 30, CAC 40, and FTSE in their local currencies between January 1, 2014 and September 20, 2015. While the Swiss stock market was leading the DAX 30, CAC 40, and FTSE 100 from January 1, 2014 through to the first half of January, the decision to remove the peg has had a noticeable impact on the performance of the SMI. Over the period January 1, 2014 to May 31, 2015, the SMI is trailing the DAX 30 by around 10 %, the CAC 40 by around 5 %, but it led the FTSE by more than 8 %. As of September 20th, the SMI is leading the DAX 30, CAC 40, and FTSE 100 by 0.5 %, 0.9 % and more than 16 %, respectively.

17.5 Review of How to Lose Money Trading Derivatives

The SNB's decision to remove the peg caused significant, and in some cases disastrous, losses at banks, hedge funds, brokerage firms, and individual traders both in and outside of Switzerland.

In this section, we discuss typical ways to lose money while trading derivatives. The underlying theme is that most disasters occur when one is not diversified in all scenarios, is overbet, and a bad scenario occurs. We can then categorize losers in the CHF black swan. Understanding how to lose helps one avoid losses!

a

b

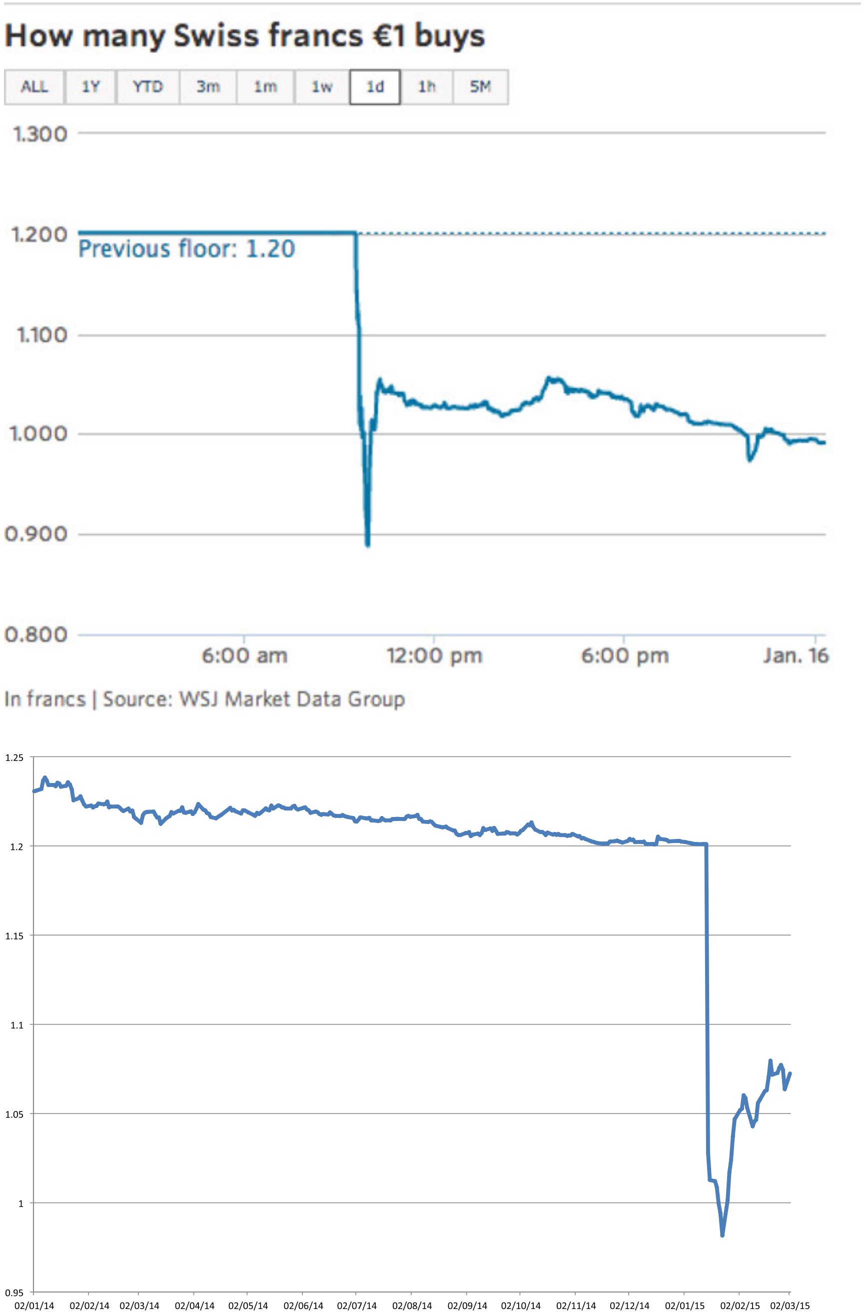
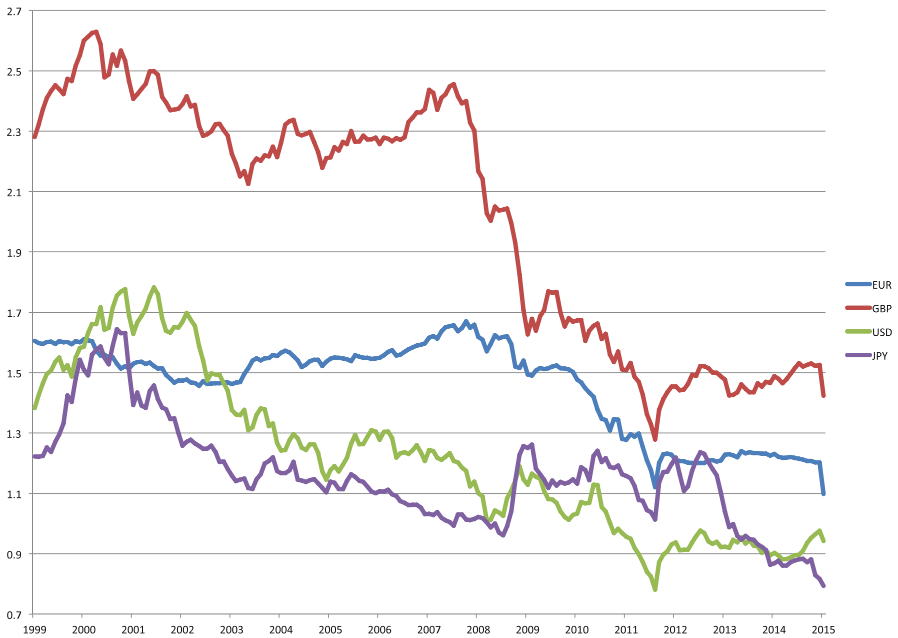


Fig. 17.3 Swiss franc U-turn. (**a**) 1-day move in the CHF/EUR exchange rate on January 15, 2015. (**b**) Evolution of the CHF/EUR from January 1, 2014 to September 21, 2015 (source: ECB)

a



b

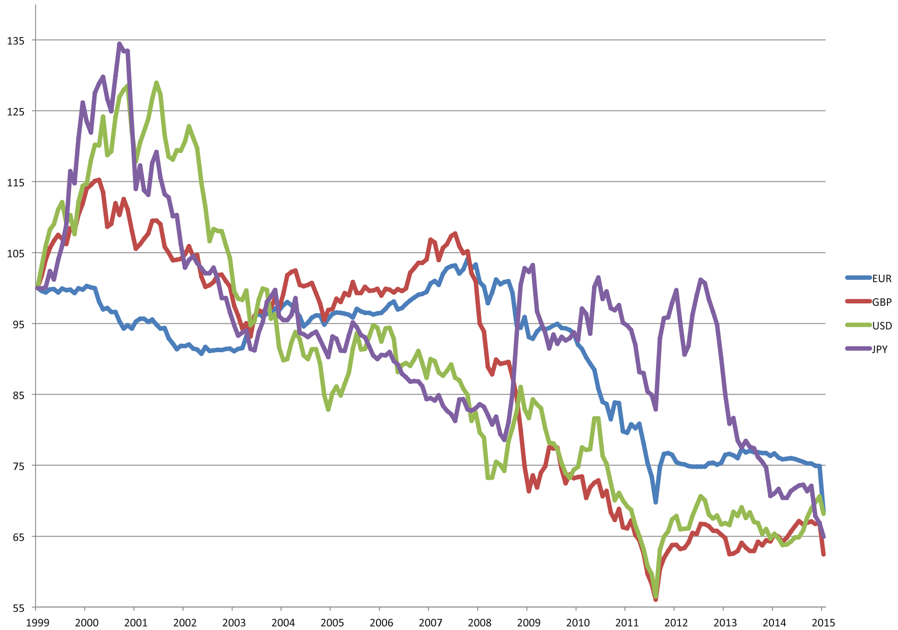


Fig. 17.4 Swiss franc U-turn: evolution of the Swiss franc against major international currencies. (**a**) Evolution of the CHF against major international currencies from January 1, 1999 to August 31, 2015 (exchange rate, source: SNB). (**b**) Evolution of the CHF against major international currencies from January 1, 1999 to August 31, 2015 (index value = 100 on January 1, 1999)

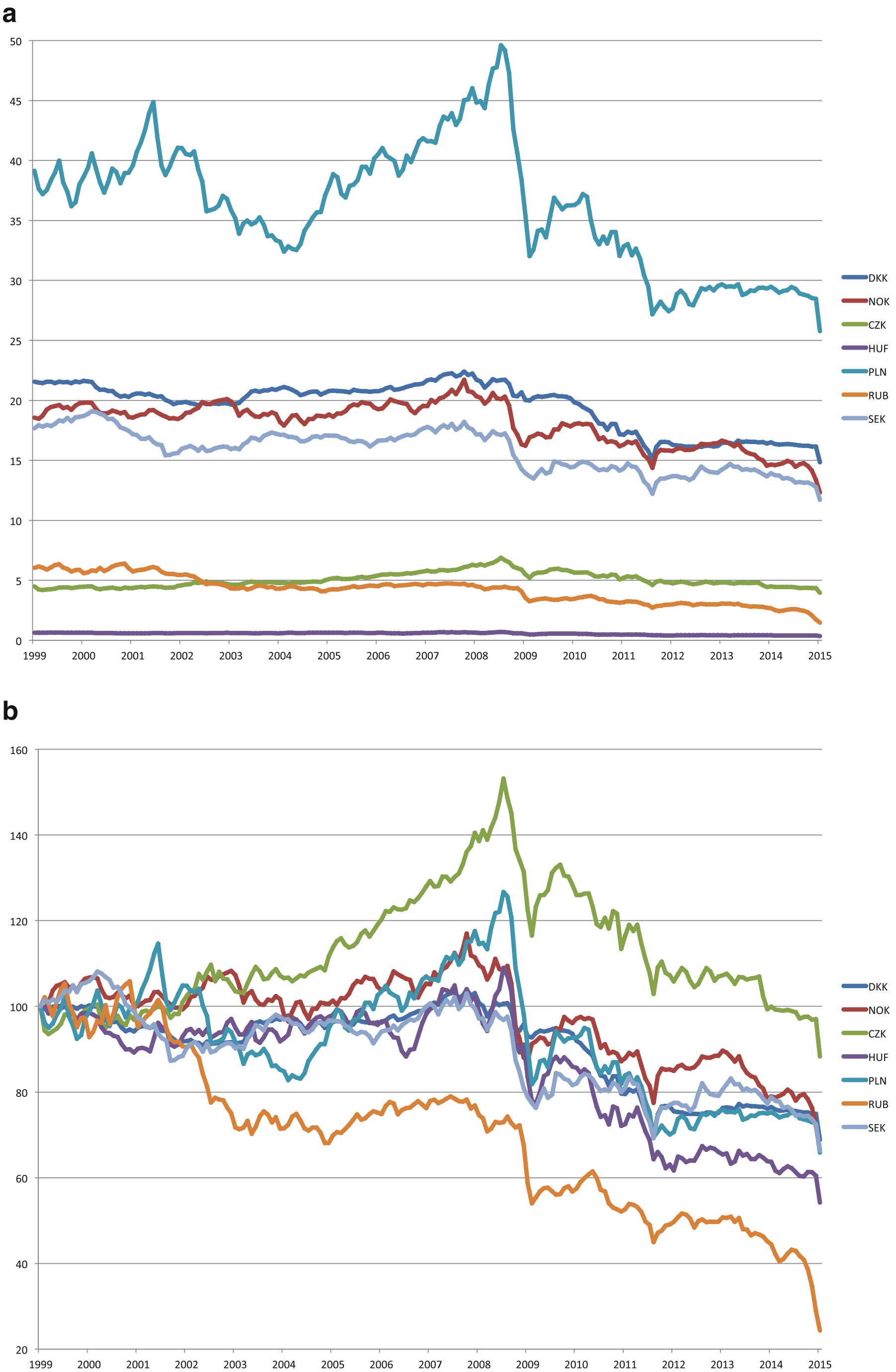
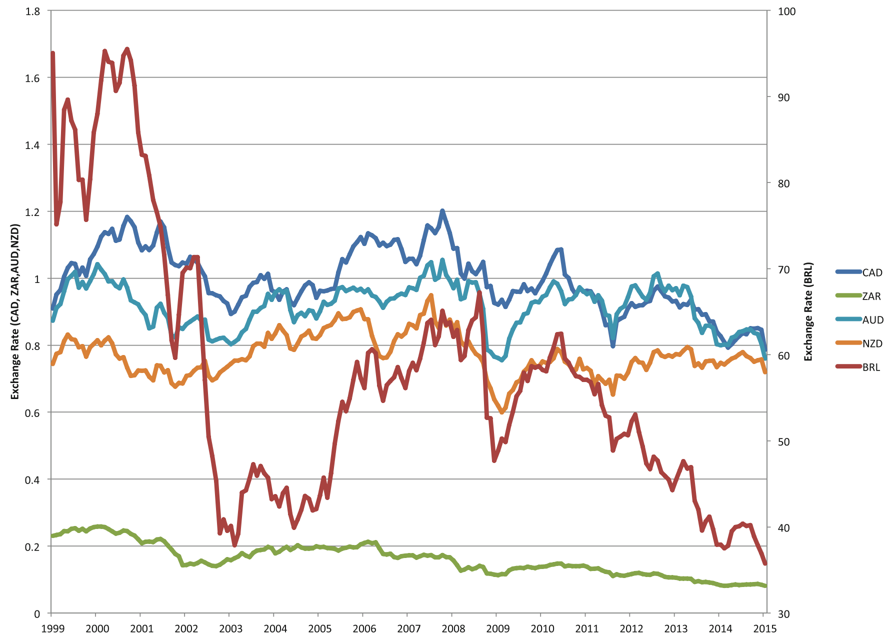


Fig. 17.5 Swiss franc U-turn: evolution of the Swiss franc against other European currencies. (**a**) Evolution of the CHF against other European currencies from January 1, 1999 to August 31, 2015 (exchange rate, source: SNB). (**b**) Evolution of the CHF against other European currencies from January 1, 1999 to August 31, 2015 (index value = 100 on January 1, 1999)

a



b

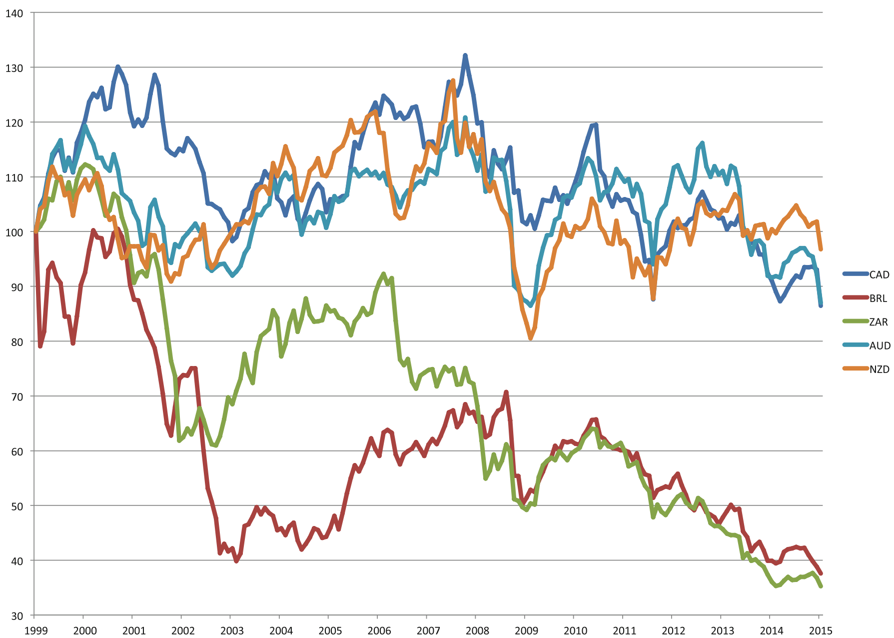


Fig. 17.6 Swiss franc U-turn: evolution of the Swiss franc against other European currencies. (**a**) Evolution of the CHF against the currencies of commodity producing countries from January 1, 1999 to August 31, 2015 (exchange rate, source: SNB). (**b**) Evolution of the CHF against the currencies of commodity producing countries from January 1, 1999 to August 31, 2015 (index value = 100 on January 1, 1999)

Fig. 17.7 Relative performance of the CHF/EUR exchange rate (*blue*), SNB's stock price (SNBN, *green*) and of the Swiss Market Index (SMI, *red*) from January 1, 2014 to September 20, 2015 (source: Yahoo! Finance) (Color figure online)

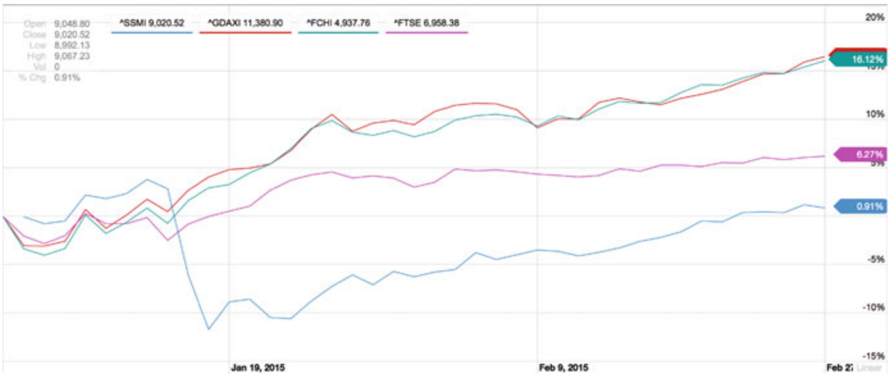


Fig. 17.8 Relative performance of the Swiss Market Index (SMI, *blue*) against the DAX 30 (*red*), CAC 40 (*green*) and FTSE 100 (*purple*) from January 1, 2014 to September 20, 2015 in their local currencies (source: Yahoo! Finance) (Color figure online)

The derivative futures industry deals with products in which one party gains what the other party loses. These are zero sum games situations. Hence there will be large winners and large losers. The size of the gains and losses is magnified by the leverage and overbetting, leading invariably to large losses when a bad scenario occurs. This industry now totals over $700 trillion of which the majority is in interest and bond derivatives with a smaller, but substantial, amount in equity derivatives. Figlewiski (1994) attempted to categorize derivative disasters and this chapter discusses and expands on that (see also Lleo and Ziemba 2015, 2014, for a discussion of banking, hedge fund and trading disasters):

1. *Hedge*

   In an ordinary hedge, one loses money on one side of the transaction in an effort to reduce risk. To evaluate the performance of a hedge one must consider all aspects of the transaction. In hedges where one delta hedges but is a net seller of options, there is volatility (gamma) risk which could lead to losses if there is a large price move up or down and the volatility rises. Also accounting problems can lead to losses if gains and losses on both sides of a derivatives hedge are recorded in the firm's financial statements at the same time.
2. *Counterparty default*

   Credit risk is the fastest growing area of derivatives and a common hedge fund strategy is to be short overpriced credit default derivatives. There are many ways to lose money on these shorts if they are not hedged correctly, even if they have a mathematical advantage. In addition, one may lose more if the counterparty defaults because of fraud or following the theft of funds, as was the case with MF Global.
3. *Speculation*

   Derivatives have many purposes including transferring risk from those who do not wish it (hedgers) to those who do (speculators). Speculators who take naked unhedged positions take the purest bet and win or lose money related to the size of the move of the underlying security. Bets on currencies, interest rates, bonds, and stock market index moves are common futures and futures options trades.

   Human agency problems frequently lead to larger losses for traders who are holding losing positions that if cashed out would lead to lost jobs or bonus. Some traders increase exposure exactly when they should reduce it in the hopes that a market turnaround will allow them to cash out with a small gain before their superiors find out about the true situation and force them to liquidate. Since the job or bonus may have already been lost, the trader's interests are in conflict with objectives of the firm and huge losses may occur. Writing options, and more generally selling volatility or insurance, which typically gain small profits most of the time but can lead to large losses, is a common vehicle for this problem because the size of the position accelerates quickly when the underlying security moves in the wrong direction as in the Victor Niederhoffer hedge fund disaster caused by the the Asian currency crisis of 1997. Since trades between large institutions frequently are not collateralized mark-to-market large paper losses can accumulate without visible signs such as a margin call. Nick Leeson's loss betting on short puts and calls on the Nikkei is one of many such examples. The Kobe earthquake was the bad scenario that bankrupted Barings.

   A proper accounting of trading success evaluates all gains and losses so that the extent of some current loss is weighed against previous gains. Derivative losses should also be compared to losses on underlying securities. For example, from January 3 to June 30, 1994, the 30-year T-bonds fell 13.6 %. Hence holders of bonds lost considerable sums as well since interest rates rose quickly and significantly.

4. *Forced liquidation at unfavorable prices*

   Gap moves through stops are one example of forced liquidation. Portfolio insurance strategies based on selling futures during the October 18, 1987 stock market crash were unable to keep up with the rapidly declining market. The futures fell 29 % that day compared to −22 % for the S&P500 cash market. Forced liquidation due to margin problems is made more difficult when others have similar positions and predicaments and this leads to contagion. The August 1998 problems of Long Term Capital Management in bond and other markets were more difficult because others had followed their lead with similar positions. When trouble arose, buyers were scarce and sellers were everywhere. Another example is Metallgellschaft's crude oil futures hedging losses of over $1.3 billion. They had long term contracts to supply oil at fixed prices for several years. These commitments were hedged with long oil futures. But when spot oil prices fell rapidly, the contracts to sell oil at high prices rose in value but did not provide current cash to cover the mark to the market futures losses. A management error led to the unwinding of the hedge near the bottom of the oil market and the disaster.

   Potential problems are greater in illiquid markets. Such positions are typically long term and liquidation must be done matching sales with available buyers. Hence, forced liquidation can lead to large bid-ask spreads. Askin Capital's failure in the bond market in 1994 was exacerbated because they held very sophisticated securities which were only traded by very few counterparties so contagion occurred. Once they learned of Askin's liquidity problems and weak bargaining position, they lowered their bids even more and were then able to gain large liquidity premiums.

5. *Misunderstanding the risk exposure*

   As derivative securities have become more complex, so has their full understanding. The (Shaw et al. 1995) Nikkei put warrant trade (discussed in Ziemba and Ziemba (2013), Chapter 12) was successful because we did a careful analysis to fairly price the securities. In many cases, losses are the result of trading in high-risk financial instruments by unsophisticated investors. Lawsuits have arisen by such investors attempting to recover some of their losses with claims that they were misled or not properly briefed on the risks of the positions taken. Since the general public and thus judges and juries find derivatives confusing and risky, even when they are used to reduce risk, such cases or their threat may be successful.

   A great risk exposure is the extreme scenario which often investors assume has zero probability when in fact they have low but positive probability. Investors are frequently unprepared for interest rate, currency or stock price changes so large and so fast that they are considered to be impossible to occur. The move of some bond interest rate spreads from 3 % a year earlier to 17 % in August/September 1998 led even savvy investors and very sophisticated Long Term Capital Management researchers and traders down this road. They had done extensive stress testing with a VaR risk model which failed as the extreme events such as the August 1998 Russian default had both the extreme low probability event plus changing correlations.

There was a similar failure of VaR and C-VaR models because of the Swiss currency unpegging, see Daníelson (2015) for discussion and some calculations. For current regulations, see Basel Committee on Banking Supervision (2013). What is needed as we argue below are convex penalty risk measures that penalize drawdowns enough to avoid the disasters. Unfortunately these types of risk functions are not yet in regulations of risk models although some applications have shown their superiority to the VaR and C-VaR models, see Geyer and Ziemba (2008), and Ziemba (2003, 2007).

Several scenario dependent correlation matrices rather then simulations around the past correlations from one correlation matrix is suggested. This is implemented, for example, in the Innovest pension plan model which does not involve levered derivative positions (see Ziemba and Ziemba 2013, Chapter 14). The key for staying out of trouble especially with highly levered positions is to fully consider the possible futures and have enough capital or access to capital to weather bad scenario storms so that any required liquidation can be done orderly.

Figlewiski (1994) observes that the risk in mortgage backed securities is especially difficult to understand. Interest only (IO) securities, which provide only the interest part of the underlying mortgage pool's payment stream, are a good example. When interest rates rise, IOs rise since payments are reduced and the stream of interest payments is larger. But when rates rise sharply, the IO falls in value like other fixed-income instruments because the future interest payments are more heavily discounted. This signal of changing interest rate exposure was one of the difficulties in Askin's losses in 1994. Similarly the sign change between stocks and bonds during stock market crashes as in 2000 to 2003 has caused other similar losses. Scenario dependent matrices are especially useful and needed in such situations.

6. *Forgetting that high returns involve high risk*

If investors seek high returns, then they will usually have some large losses. The Kelly criterion strategy and its variants provide a theory to achieve very high long term returns but large losses will also occur. These losses are magnified with derivative securities and especially with large derivative positions relative to the investor's available capital.

7. *How over betting occurs*

Figure 17.9 shows how the typical over bet situation occurs assuming a Kelly strategy is being used. The top of the growth rate curve is at the full Kelly bet level that's the asset allocation maximizing the expected value of the log of the final wealth subject to the constraints of the model. To the left of this point are the fractional Kelly strategies which under a lognormal asset distribution assumption use a negative power utility function rather than log. So αw^{α}, for $\alpha < 0$ gives the fractional Kelly weight $f = \frac{1}{1-\alpha}$. So $u(w) = \frac{-1}{w}$ corresponds to $\frac{1}{2}$ Kelly with $\alpha = -1$. Overbetting is to the right of the full Kelly strategy and it is clear that betting more than full Kelly gives more risk measured by the probability of reaching a high goal before a lower level curve on the figure. It is this area way to the right where over betting occurs. And virtually all of the disasters occur because of the over betting.

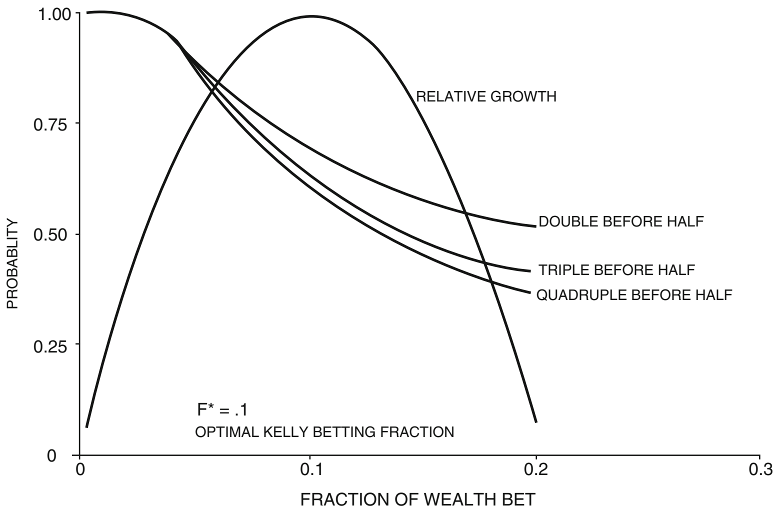


Fig. 17.9 Relative growth and probabilities of doubling, tripling, and quadrupling initial wealth for various fractions of wealth bet for the gamble win \$2 with probability 0.4 and lose \$1 with probability 0.6

It is easy to over bet with derivative positions as the size depends on the volatility and other parameters and is always changing. So a position safe one day can become very risky very fast. A full treatment of the pros and cons of Kelly betting is in Ziemba (2015).

Stochastic programming models provide a good way to try to avoid problems 1–6 by carefully modeling the situation at hand and considering the possible economic futures in an organized way.

Hedge fund and bank trading disasters usually occur because traders overbet, the portfolio is not truly diversified and then trouble arises when a bad scenario occurs. Lleo and Ziemba (2015) discuss a number of sensational failures including Metalgesllshart (1993), LTCM (1998), Niederhoffer (1997), Amaranth Advisors (2006), Merrill Lynch (2007), Société Généralé (2008), Lehman (2008), AIG (2008), Citigroup (2008), MF Global (2012), and Monte Paschi (2013). Stochastic programming models provide a way to deal with the risk control of such portfolios using an overall approach to position size, taking into account various possible scenarios that may be beyond the range of previous historical data. Since correlations are scenario dependent, this approach is useful to model the overall position size. The model will not allow the hedge fund to maintain positions so large and so under diversified that a major disaster can occur. Also the model will force consideration of how the fund will attempt to deal with the bad scenario because once there is a derivative disaster, it is very difficult to resolve the problem. More cash is immediately needed and there are liquidity and other considerations. Ziemba and Ziemba (2013), Chapter 14 explores more deeply such models in the context of pension fund as well as hedge fund management.

Litzenberger and Modest (2009), who were on the firing line for the LTCM failure, propose a modification of standard finance CAPM type theory modified for fat tails and C-VaR or expected tail losses for the losses. Ziemba (2003, 2007, 2013) presents his approach using convex risk measures and three scenario dependent correlation matrices depending upon volatility using stochastic programming scenario optimization. Both of these approaches would mitigate such losses. The key is not to over bet and have access to capital once a crisis occurs and to plan in advance for such events.

17.6 The Folly of the Misleading Value at Risk Measure

Value at risk (VaR) is the most widely used risk measure and has held a central place in the development of international banking regulations in general and of the Basel Accord in particular. The VaR of a portfolio represents the maximum loss within a confidence level of $1 - \alpha$ (with a between 0 and 1) that the portfolio could incur over a specified time period, for instance a d-days horizon (see Fig. 17.10). For example, if the 10-day 95 % VaR of a portfolio is \$10 million, then the expectation with 95 % confidence is that the portfolio will not lose more than \$10 million during any 10-day period. The (1 - α) VaR of a portfolio with (random) P&L X is defined as

$$\mathrm{VaR}(X;\alpha) = -\{X|F(X) \leq \alpha\},$$

which reads "minus the loss X (so the VaR is a positive number) chosen such that a greater loss than X occurs in no more than a percent of cases."

Jorion (2006) presents a comprehensive and highly readable reference on VaR and its use in the banking industry, while Embrechts et al. (2005) cover risk management from a mathematical and technical perspective. Value at Risk has the advantage of being a particularly simple risk measure, because it corresponds to minus the α-quantile of the P&L distribution:

$$\mathrm{VaR}(X;\alpha) = -q_\alpha(X).$$

VaR is also elicitable (see Ziegel 2014), a property shared by all the quantiles.

An alternative definition for the VaR of a portfolio is the *minimum* amount that a portfolio is expected to lose within a specified time period and at a given confidence level of a reveals a crucial weakness. The VaR has a well-documented "blind spot" in the α-tail of the distribution, which means that it is impossible to evaluate the probability and severity of truly extreme events. The P&L distributions for investments X and Y in Fig. 17.11 have the same VaR, but the P&L distribution of Y is riskier because it has larger potential losses.

Artzner et al. (1999) defined coherent risk measures as the class of monetary risk measures satisfying four "coherence" axioms. VaR is not a coherent risk measure.

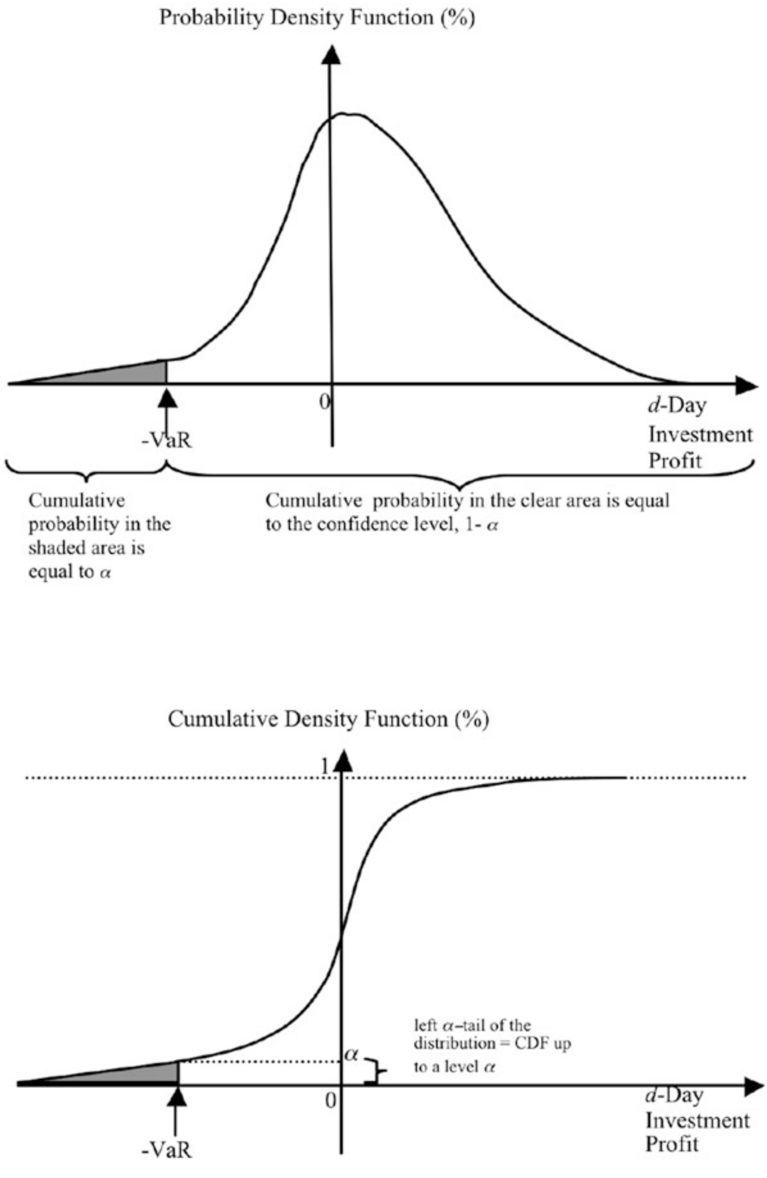


Fig. 17.10 Value-at-Risk in terms of both PDF and CDF

Monetary risk measures, introduced by Artzner et al. (1999), is a class of risk measures that equate the risk of an investment with the minimum amount of cash, or capital, that one needs to add to a specific risky investment to make its risk acceptable to the investor or regulator. A monetary measure of risk r is defined as

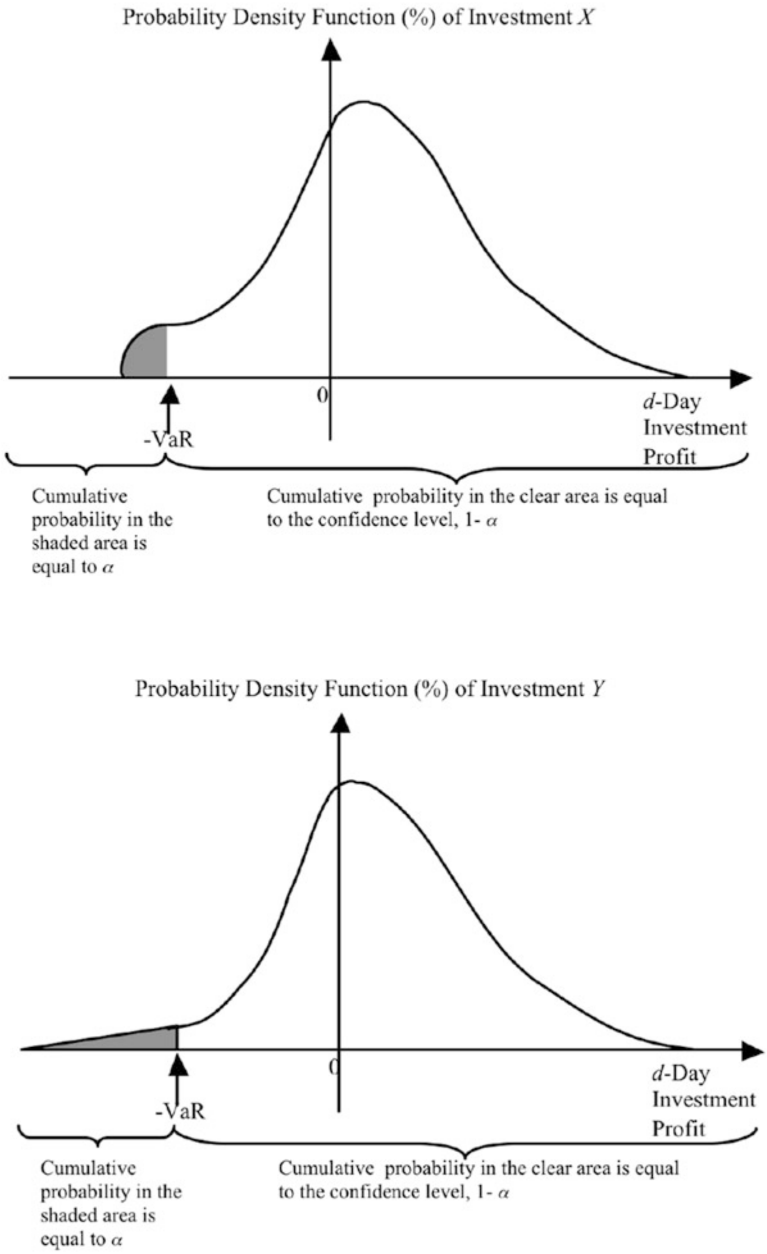


Fig. 17.11 Two investments with same Value-at-Risk, but different degrees of desirability

$$\rho(X) := \min_{k \geq 0} [\text{an investment in a position } (X + k) \text{ is acceptable}],$$

where k represents an amount of cash or capital and X is the monetary profit and loss (P&L) of some investment or portfolio during a given time horizon, and discounted back to the initial time.

The coherence axioms are

1. Monotonicity: if the return of asset X is always less than that of asset Y, then the risk of asset X must be greater. This translates into

$$X \leq Y \text{ in all states of the world} \Rightarrow \rho(X) \geq \rho(Y). \tag{A1}$$

2. Subadditivity: the risk of a portfolio of assets cannot be more than the sum of the risks of the individual positions. Formally, if an investor has two positions in investments X and Y, then

$$\rho(X + Y) \leq \rho(X) + \rho(Y). \tag{A2}$$

   This property guarantees that the risk of a portfolio cannot be more (and should generally be less) than the sum of the risks of its positions, and hence it can be viewed as an extension of the concept of diversification introduced by Markowitz. This property is important for portfolio managers and banks trying to aggregate their risks among several trading desks. VaR is not subadditive, so VaR may not reward diversification, which potentially results in increased concentration risk.
3. Homogeneity: if a position in asset X is increased by some proportion k, then the risk of the position increases by the same proportion k. Mathematically,

$$\rho(kX) = k\rho(X). \tag{A3}$$

   This property guarantees that risk scales according to the size of the positions taken. This property, however, does not reflect the increased liquidity risk that may arise when a position increases. For example, owning 500,000 shares of company XYZ might be riskier than owning 100 shares because in the event of a crisis, selling 500,000 shares will be more difficult, costly, and require more time. As a remedy, Artzner, Delbaen, Eber, and Heath proposed to adjust X directly to reflect the increased liquidity risk of a larger position.
4. Translation invariance or risk-free condition: adding cash to an existing position reduces the risk of the position by an equivalent amount. For an investment with value X and an amount of cash r,

$$\rho(X + r) = \rho(X) - r. \tag{A4}$$

Stress testing complements VaR by helping address the blind spot in the α-tail of the distribution. In stress testing, the risk manager analyzes the behavior of the portfolio under a number of extreme market scenarios that may include historical scenarios as well as scenarios designed by the risk manager. The choice of scenarios and the ability to fully price the portfolio in each situation are critical to the success of stress testing. Jorion (2006) discussed stress testing and how it complements VaR.

Conditional VaR (CVaR) is an improvement over VaR. Conditional VaR is the average of all the d-day losses exceeding the d-day $(1-\alpha)$ VaR (see Fig. 17.12). Thus, the CVaR cannot be less than the VaR, and the computation of the d-day $(1-\alpha)$ VaR is embedded in the calculation of the d-day $(1-\alpha)$ CVaR.

Formally, the d-day $(1-\alpha)$ CVaR of an asset or portfolio X is defined as

$$\mathrm{CVaR}(X;\alpha) = -E\left[X|X \le F_X^{-1}(\alpha)\right]. \tag{A5}$$

This formula takes the inverse CDF of the confidence level, a, to give a monetary loss threshold (equal to the VaR). The CVaR is then obtained by taking the expectation, or mean value of all the possible losses in the left tail of the distribution, beyond the threshold. CVaR is a coherent risk measure, implying that it accounts for diversification, and it can be used efficiently to optimize portfolios (see Rockafellar and Uryasev 2000, 2002).

However, CVaR is not elicitable, depends heavily on the quality of tail data, and only introduces a linear penalization for the loss. This last point comes from the definition CVaR as the mean tail loss operator, implying that CVaR computes risk as a linear function of tail loss.

As an alternative, Ziemba (2013) has argued for convex risk measures that penalize losses more and more as the losses mount. Rockafellar and Ziemba (2000, 2013) define convex risk measures as monetary risk measures satisfying the five following axioms:

$$\rho(X+\alpha\cdot r) = \rho(X)-\alpha. \tag{R1}$$

$$\rho(\lambda X+(1-\lambda)\,Y) \le \lambda\rho(X)+(1-\lambda)\,\rho(Y), 0\le\lambda\le 1. \tag{R2}$$

$$X\le Y \Rightarrow \rho(X)\ge\rho(Y). \tag{R3}$$

$$X<0 \Rightarrow \rho(X)>0. \tag{R4}$$

$$\rho(0)=0. \tag{R5}$$

In Rockafellar and Ziemba's definition, axioms (R2) and (R5) replace the more restrictive coherence axioms (A2) and (A3). Separately, Föllmer and Schied (2002) proposed an alternate definition of convex risk measure simply replacing coherence axioms (A2) and (A3) by the convexity property (R2).

However, the industry still uses the flawed value at risk which penalizes a loss of 1 billion the same as 1 million if 1 million is the VaR number to be exceeded only 5 % of the time. Shorting the CHF was a popular trade and most firms would lever their position some 20 times or more. With such leverage a 5 % move against the position wipes out all the value. Yet the trades were seen as relatively low risk using VaR models at financial institutions because volatility of the CHF was reduced by

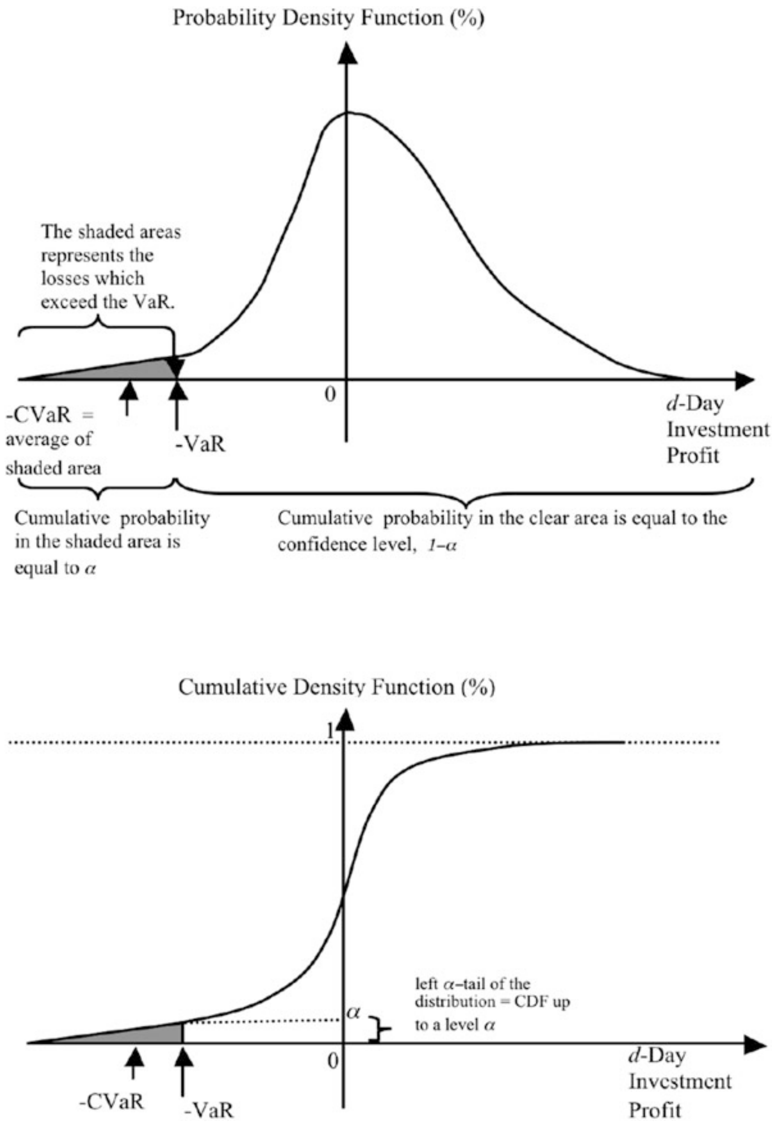


Fig. 17.12 Conditional Value-at-Risk in terms of both PDF and CDF

the SNB's cap. See Daníelson (2015) for an analysis of the failure of VaR and C-Var risk models around the time of the Swiss currency unpegging. An important point here is that the regulations do not include the more realistic convex risk measures.

Regardless of the risk measure, the design of scenarios and stress tests is crucial. Ziemba (2003, 2007) makes the following remark and establishes a list of factors that should be considered when designing scenarios:

> The generation of good scenarios that well represent the future evolution of the key parameters is crucial to the success of the modeling effort. Scenario generation, sampling, and aggregation are complex subjects, and I will discuss them by describing key elements and providing various developed and implemented models.
> Scenarios should consider:
>
> - mean reversion of asset prices;
> - volatility clumping, in which a period of high volatility is followed by another period of high volatility;
> - volatility increases when prices fall and decreases when they rise; trending of currency, interest rates, and bond prices;
> - ways to estimate mean returns;
> - ways to estimate fat tails; and
> - ways to eliminate arbitrage opportunities or minimize their effects.

Depending on the specificities of the problem, we could use either of the following five methods to generate scenarios:

1. full knowledge of the exact probability distribution, *P*;
2. use a known parametric family of statistical or probabilistic models;
3. moment matching;
4. historical simulation;
5. expert opinion

We could also combine several methods. For example, the Black–Litterman model (see Black and Litterman, 1992) and its descendants combine a parametric approach with expert opinions.

17.7 Losers and How It Affected Them

The major economic activity in Switzerland is money storage and management. They produce watches, chocolate, pharmaceuticals and tourist activities such as skiing, hiking, and visiting the beautiful countryside. Half the GDP comes from exports. Dhubat (2015) discusses the Canadian-Swiss chocolate market. The CAD went from 0.85 to 0.74 CHF with the unpegging, some 13.75 % more expensive. When Ziemba sold a VW camper bus in Zurich in 1973 each CAD was worth 4 CHF. Ziemba got CHF12,000 for a six-month-old camper that cost C$3,000 when purchased. This CHF 12,000 is now worth C$15,751.60 after the revaluation. This reminds us that the CHF has changed dramatically overtime. Canada has about US$2.7 billion in chocolate sales versus total North American sales of about US$20.2. Toronto-based Swiss national chocolatier Ingrid Laderach who sells mostly Swiss made chocolate: “it was a huge shock, and being Swiss myself, I was kind of disappointed at my countrymen to be honest with you.” She expects that the SNB’s move would impact Valentine’s day and Easter when chocolate demand is high. This of course is a plus for local Canadian producers.

Ski resorts in Switzerland have been hit hard as their prices are about double those in France and Austria. So Swiss resorts have needed to lower their prices—especially for foreigners.

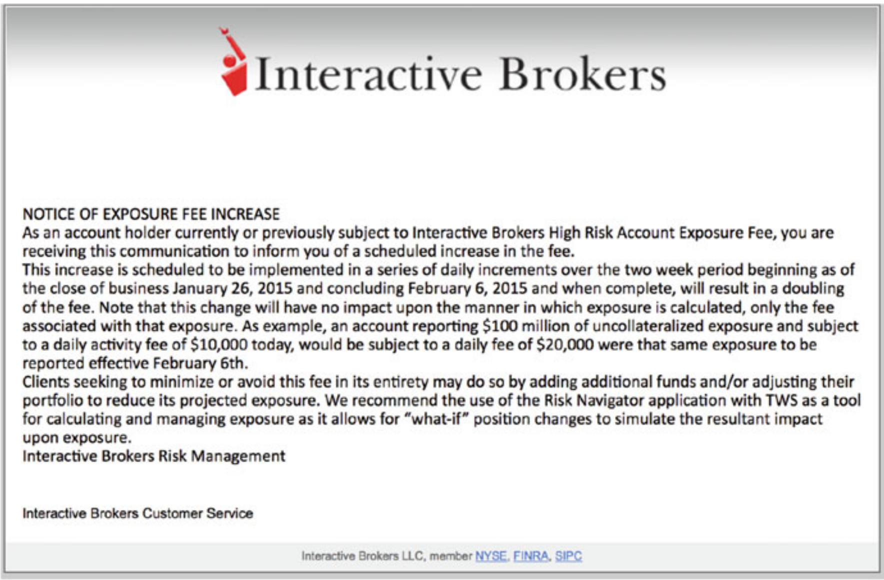
Swiss research institutes have lowered growth forecasts by 75 %. Companies are squeezing employees with lower pay and more hours of work. Retailers are cutting prices and have the added problem of cross border shopping into France, Germany, and Italy. High end products such as the top watches made by TAG Heuer and others are less hit as their very high profit margins act as a buffer against currency shocks. Private wealth management banks face significant difficulties because they are forced to be more transparent. In addition their traditional competitive advantages, such as secrecy and a perception of safety are declining.

17.8 Banks and Hedge Funds

The losses are in the billions: Citigroup Inc, Deutche Bank AG, and Barclays PLC together lost US$400 million. Marco Dimitrojevic's US$830 million hedge fund was hit so bad that it had to be closed.

Interactive Brokers (IB) is a web-based brokerage firm that is growing rapidly, offering attractive terms to traders. IB has been rated #1 by *Barron's* 3 years in a row, is a stock pick of Motley Fool, and is highly regarded. They have low fees for electronic trading. They are aware of possible losses and do certain things to prevent them. They reported that several customers suffered losses in excess of their account capital, amounting to about $120 million which is about 2.5 % of the net worth of the company. Ziemba has accounts with them: what they do is charge an insurance fee if you have positions such that a 30 % up or down move would wipe out all your capital. It is not clear whether they buy the puts for this or simply pocket the money as part of their business (a form of self-insurance). Possibly because of the 120 million loss they are doubling the exposure fee, see Fig. 17.13.

Very big losers were small time individual retail FOREX traders and the firms they traded with. These individuals expect to win but in fact most lost because of the volatility of the market and the fact that they are undercapitalized (or in other words, over levered). The Aite Group LLC found that 11 % of such traders expect to lose while the other 89 % expect to win, fully 41 % expect to gain 10 % per month. Citi estimated worldwide that there are some 4 million such traders with about 150,000 in the USA. The NFA estimated that 72 % lose money. Alpari which folded on Friday after the January 15 2015 unpegging had about 70,000 such clients. Gain Capital was growing customer trading volume at 90 % per year and their income was growing even faster from US$7 million in 2004 to US$230 million in 2008. The firm was allowing huge leverage, for example, a cash account with $5,000 could control $1 million in currency positions which is 200:1 leverage. Some firms, like Gain, take the other side of the trades, so these small traders were not client customers but simply counterparties. Gain, of course and others like them, won. With such leverage and high volatility, most clients are losers. In the USA the NFA required large capital



Interactive Brokers

NOTICE OF EXPOSURE FEE INCREASE
As an account holder currently or previously subject to Interactive Brokers High Risk Account Exposure Fee, you are receiving this communication to inform you of a scheduled increase in the fee.
This increase is scheduled to be implemented in a series of daily increments over the two week period beginning as of the close of business January 26, 2015 and concluding February 6, 2015 and when complete, will result in a doubling of the fee. Note that this change will have no impact upon the manner in which exposure is calculated, only the fee associated with that exposure. As example, an account reporting $100 million of uncollateralized exposure and subject to a daily activity fee of $10,000 today, would be subject to a daily fee of $20,000 were that same exposure to be reported effective February 6th.
Clients seeking to minimize or avoid this fee in its entirety may do so by adding additional funds and/or adjusting their portfolio to reduce its projected exposure. We recommend the use of the Risk Navigator application with TWS as a tool for calculating and managing exposure as it allows for "what-if" position changes to simulate the resultant impact upon exposure.
Interactive Brokers Risk Management

Interactive Brokers Customer Service

Interactive Brokers LLC, member NYSE, FINRA, SIPC

Fig. 17.13 Interactive Brokers portfolio insurance charges for risky positions in the USA

and permitted a 50:1 leverage so much of the business moved to London where leverage up to 500:1 was allowed and in Cyprus 1000:1 leverage was possible. To get more customers, such firms have extensive marketing because old customers are blowing up and leaving, but there is always another sucker out there for them. Spot FOREX trading is not regulated in London or Europe. In London, the financial conduct authority (FCA) only takes action if there is fraud or boiler rooms in action (of course, one could suggest that this entire segment of the industry is fraudulent).

A prime example with a different type of operating procedure was Drew Nir's firm FXGM as reported by Evans (2015) and Lex team (2015). FXGM has a 157 page prospectus which has one dangerous provision for themselves: they do not try to obtain more funds or sue clients who lose money or go into negative equity. They allowed 29 % of their clients to use credit cards even though that is not allowed and they are not on the other side but they are supposed to hedge. Their clients lost about $225 million. FXGM are allowed forced sale of customer positions in deficit but in this case the currency move was way too fast to do much of this. Before January 15, they had a market cap of $1.4 billion and $300 million capital, which was $200 million above the $100 million required by the regulators. Its shares had been listed in 2010 at $14. They handled $1.4 trillion of trades in Q4:2014. Post peg the stock in FXGM fell 87.33 % to $1.60 and as low as 98 cents from a pre-January 15 price of about $12.63. They needed a bailout and after the market closed on Friday, Jeffries arranged a $300 million loan at 10 % interest from Leucadia National Corp. The shares rebounded and were worth $2.43 on January 23, 2015.

Another problem is co-mingled funds which is always a big danger. Stock and future funds do not legally allow co-mingling of client and firm funds. MF Global (see Lleo and Ziemba 2015, 2014) is one example where this policy was violated. In FXGM's case, this inadvertent co-mingling led to their large losses.

17.9 What Types of Traders Lost Money

Actually one did not have to be doing trades to lose money. Anyone in Switzerland holding foreign currencies took losses if they want to spend their money at home.

One Swiss colleague who is a professor and trader had the trading firm's capital in US$ since many of their trades are in US$. Another professor colleague who is German but at a university in Switzerland plans a retirement in Germany so his CHF holdings including his university pension gained in EUR terms but his other assets in other currencies loss in CHF terms, he writes that he gained 10 % in EUR and lost 10 % in CHF in his overall situation.

Some traders that lost money:

1. Short puts on the EUR/CHF cross. One bets that the EUR will not fall and collects a small premium. The tails here follow typical deep out of the money favorite-longshot bias characteristics. Ziegler and Ziemba (2015) study returns from buying and selling hedged and unhedged puts and calls from 1985 to 2010 on the S&P 500 futures. These types of trades usually win but if there is a big move in the wrong direction, the losses can be very large. The Niederhoffer bankruptcy from the Asian currency crisis in 1997 is a typical example (Ziemba and Ziemba 2013). $120 million in his hedge fund was turned into $70 million by buying cheap Thai stocks which continued to drop. Then the $70 million was turned into −$20 million by shorting out of the money S&P puts. It turned out that the puts expired worthless the next month and Nieferhoffer would have survived if he had more capital. This is another reminder that one needs to be sufficiently capitalized for the type of trading one is doing. Usually one must have a large capital being each short position to try to weather storms. In the CHF case, a 15 % move yielded large losses. All this depends on how fast the brokerage firm is checking the positions. Minutes after the announcement the CHF was up 38 % and then settled at 15 % ahead. Of course, those who were long CHF with short calls on the EUR/CHF would gain but only the premium. So the losses are much greater than the gains in this case.
2. Short strangles and straddles: these involve selling both sides of the market, that is short puts and short calls, collecting two premiums. One has a strike at the money and the other is out of the money. Those like in (1) would have large losses less the two premiums which would not lower the loss much. The opposite position, buying the puts and calls, paying two premiums, which is usually a losing strategy, in this case would have had had huge gains on the long CHF side.

17.10 Mortgage Losses

Swiss fixed interest rates on mortgages are as low as 1.5 %. With the exchange rate fixed, borrowing in Swiss for homeowners in Austria, Hungary, Poland, and other countries in central and eastern Europe seemed like a good decision.[3] During the real estate bubble of 2005–2007, mortgage rates in these countries were over 10 %. In addition, the local currencies were rising in value as investors anticipated Poland and Hungary joining the eurozone. But in 2008, these currencies fell relative to the strong Swiss franc so the payments increased. Poland banned new CHF lending and Hungary added the increased payments to the principal owing.

There was a disconnect between western Europe and eastern Europe Swiss franc borrowers. Those in the west are concentrated in the business and financial sectors and more care was taken to hedge the currency risk. But in eastern Europe, it was mortgages, some 566,000 Polish, 150,000 Romanian and 60,000 Croatian. And in Hungary, half of all households in the country had foreign currency debt with most in Swiss francs. The mortgages are not transparent as mostly the interest is paid in local currency although it is computed in Swiss francs at a non-Swiss bank.

In 2011 the same flight to safety that devalued the Polish zloty and Hungarian forint in 2008 devalued the euro against the Swiss franc so all costs in francs increased such as these loans.

Croatia pegged its currency, the kuna to the franc for one year. This had a large cost of at least 30 % of its currency reserves. And if the franc continues to rise against the euro, Crotatian goods will be more expensive for French, German, Italian and other customers. Other countries such as Romania are considering similar moves. In contrast, Hungary reacted by forcing all mortgages such as these to convert their Swiss franc loans to Hungarian forints. Foreign banks, and in particular Austrian banks, had to bear the adjustment cost.

17.11 Final Remarks

The Quantitative Easing programme implemented by the US Federal Reserve as a remedy for the 2007–9 financial crisis has led to a massive increase in US stock prices, tripling the S&P 500 index since the March 2009 low. Unemployment is now much lower even though wages have not increased. This sets the stage for gradual interest rate hike, limited in scope by the huge debt loads of the US government. With its USD 4 trillion balance sheet, the Federal reserve Bank has a tricky policy road ahead and there likely will be some bumps along the way. We see this already with daily up and down moves in the US stock market as the probability of when the first rate rise will occur is reassessed daily.

[3]This section is based on Frum (2015) and other sources.

Now, the European Central Bank is embarking on the same path. The money they spend to buy bonds does not directly go to the real problem, namely unemployment. But rather it will basically go to the banks and lower interest rates. The divergence between US rates heading up and European rates going down has helped fuel a massive shift to the US dollar with most countries currencies dropping. The Swiss franc has historically been a safe haven and it has been for many years on a monotone increase in value. Jim Rogers noted this as did Ziemba's car sale in the 1970s.

Since then the franc has increased fivefold and is continuing to strengthen. Although many aspects of the Swiss banking advantage are declining, there is still much demand for the currency even with negative interest rates. This paper concerns the January 15, 2015 unpegging on the Euro Franc exchange rate at 1.20. Since 2011, the Swiss central bank, partially owned by the cantons, had been buying various currencies to keep the euro exchange rate at this level. This caused them considerable losses. These losses were a factor in their decision to exit the peg. We discussed this abrupt action that caused a large, fast move in currency prices and triggered large losses for many individuals and institutions.

After the markets calmed down the currencies were roughly 15 % lower. The Swiss National Bank subsequently announced a loose target zone of 1.05–1.10 for the Euro. They were prepared to spend considerable funds to keep this range. Meanwhile the Euro has declined steeply against the US Dollar. Even short covering rallies are met with more selling. The Euro, which once fetched 1.60 US dollars and was 1.40 in May 2014, was down to 1.05 when we went to press.

There are conflicting views of the situation. Some forecasters, such as Deutsche Bank, predicted a move to 1.00 by the end of 2015 and a new cycle low of 85 cents by 2017. Their reasoning articulates around the concept of a "Euro glut." Deutsche Bank explains: "simply put, it argues that the euro-area's gigantic current account surplus, combined with the European Central Bank's Quantitative Easing program, and with negative interest rates will continue to cause the Euro to tumble?" Robin Winkler and George Saravelos of the Deutsche Bank say that the region, which is currently a debtor to the world, must become a net creditor to the word. To that end, its investment position needing to reach 30 % of GDP versus its current −10 % before the current account surplus is sustainable. This can only happen with net capital outflows of at least 4 trillion Euros. In fact, European outflows in the last six months have been high, putting downward pressure on the Euro/US dollar exchange rate.

Rachel Ziemba disagrees with this and argues that the EuroZone has net foreign assets (stock) and net surplus (flow). Until the Euro crisis, Europe's balance of payments was roughly balanced since the net surplus from the core countries especially Germany offsets deficits in the periphery countries. Currently these latter countries are doing (a) fiscal adjustment, (b) structural reforms, and (c) deleveraging to some extent. Hence the deficits have shrunk even though the stock of debt is high. She argues that the Deutsche bank conflates two drivers of Euro weakness, namely the Quantitative easing and lower rate differential versus the US with one (current account surplus) that is supportive of Euro strength. Should the trade and capital flows reverse because of these ECB actions, the Euro weakness could extend, but the weaker Euro lowers imports and increases exports.

So what does this mean for the Swiss Franc? For Switzerland as for the USA, it certainly seems that a continued increase in the respective values is likely until the exchange rate greatly affects the economy of Switzerland and the USA. As we go to press in September 2015, the Euro/Franc exchange rare is near the bottom of its target range about 1.08 and the US dollar is close to par with the Swiss Franc with a 0.9719 exchange rate.

Acknowledgements The first author "Sebastien Lleo" gratefully acknowledges the support from Région Champagne Ardennes and the European Union through the RiskPerform grant.

References

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1997, November). Thinking coherently. *Risk*, *10*, 68–71.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999, July). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203–228.

Barker, B. (2015). What happened in Switzerland?

Basel Committee on Banking Supervision (2013). Fundamental review of the trading book: A revised market risk framework. Technical Report, Bank for International Settlements: Basel, Switzerland.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, *48*, 28–43.

Chupka, K., & Schiff, Q. E. (2015). *Won't work in Europe just like It didn't work here*. Available online: http://finance.yahoo.com/news/schiff--qe-won-t-work-in-europe-just-like-it-didn-t-work-here-155615061.html. Accessed on 27 Jan 2015.

Daníelson, J. (2015). *What the Swiss FX shock says about risk models*. CEPR's Policy Portal. Available online: http://www.voxeu.org/article/what-swiss-fx-shock-says-about-risk-models. Accessed on 1 Mar 2015.

Davies, G. (2015). The Swiss currency bombshell - cause and effect. Financial Times, January 18.

Dhubat, H. (2015). Swiss franc shock bitter for chocolatiers. Globe and Mail, January 21.

Economist (2015a). Better late than never. January 24.

Economist (2015b). Broke but never bust. January 24.

Economist (2015c). Swiss miss. January 24.

Economist (2015d). Switzerland's currency: shaken not stirred. January 24.

Economist (2015e). Why investors would opt to lose money. January 24.

Embrechts, P., Frey, R., & McNeil, A. (2005). *Quantitative risk management: concepts, techniques and tools*. Princeton, NJ, USA: Princeton University Press.

Evans, D. (2015) *How Swiss shock humbled king of leveraged currency trading*. Available online: http://www.bloomberg.com/news/articles/2015-01-20/how-swiss-shock-humbled-the-king-of-leveraged-currency-trading. Accessed on 15 Apr 2015.

Figlewiski, S. (1994). How to lose money in derivatives. *Journal of Derivatives* (Winter), 75–82.

Financial Times (2015). A poor advertisement for Swiss reliability. January 15.

Föllmer, H., & Schied, A. (2002). *Advances in finance and stochastics*, Chapter Robust representation of convex measures of risk. Berlin Heidelberg, Germany: Springer.

Frum, D. (2015). The world's next mortgage crisis? *The Atlantic*.

Gauriat, V. (2015). Businesses struggle to adapt to strong Swiss franc. Special Reports, Euronews, Feb 20.

Geyer, A., & Ziemba, W.T. (2008). The Innovest Austrian pension fund planning model InnoALM. *Operations Research*, *56*, 797–810.

Jorion, P. (2006). *Value at risk: The new benchmark for managing financial risk* (3 ed.). New York, NY, USA: McGraw-Hill.

Krugman, P. (2015a). Francs, fear and folly. *New York Times*, January 15.

Krugman, P. (2015b). A tale of two pegs. The New York Times, January 21.

Lexteam (2015). Currency brokers: special fx. January 16.

Litzenberger, R., & Modest, D. (2009, April). Crisis and non-crisis risk in financial markets: A unified approach to risk management. Technical report, Working Paper.

Lleo, S., & Ziemba, W. (2015). How to lose money in derivatives: Examples from hedge funds and bank trading departments. In A. Malliaris, & W. Ziemba (Eds.), *Handbook of futures markets*. Singapore, Singapore: World Scientific Publishing Co.

Lleo, S., & Ziemba, W. T. (2014). How to lose money in the finacial markets: examples from the recent financial crisis. *Alternative Investment Analyst Review*, *3*(3), 22–35.

MacLean, L., & Ziemba, W. (2013). *Handbook of the fundamentals of financial decision making*. World Scientific Handbook in Financial Economics. Singapore, Singapore: World Scientific Publishing Company Pvt. Limited.

Malpass, M. (2015). Monetary dead end: the EB embraces quantitative easing despite the sorry track record of "helicopter money'. *Financial Post*, January.

Minder, R., & Granelsky, V. (2015). Swiss watchmakers jolted by currency shock, January 22.

Murphy, P. (2015). Hapless horde of retail investors take hit in Swiss havoc. Financial Times, January 16.

Phillips, M. (2015). Absolutely everything you need to understand what happened to the Swiss franc this week. January 18.

Rockafellar, R. T., & Ziemba, W. (2000). Modified risk measures and acceptance sets. Technical Report Mimeo, University of Washington, July.

Rockafellar, R. T., & Ziemba, W. (2013). Modified risk measures and acceptance sets. In L. MacLean, & W. Ziemba (Eds.), *Handbook of the fundamentals of financial decision making: Part I*, Chapter 29, pp. 505–506. July: World Scientific Publishing Co.

Rockafellar, T., & Uryasev, S. (2000, Spring). Optimization of conditional value-at-risk. *Journal of Risk*, *2*(3), 21–41.

Rockafellar, T., & Uryasev, S. (2002, July). Conditional value-at-risk for general loss distribution. *Journal of Banking & Finance*, *26*(7), 1443–1471.

Rogers, J. (2013). *Street smarts*. New York: Crown Business.

Rogers, J. (2015). I warned you the Swiss Central Bank's currency policy would end disastrously. Business Insider, January 16.

Schwartz, S. (2015). *Independent economist*. Private conversation. Canada: Vancouver.

Shaw, J., Thorp, E. O., & Ziemba, W. T. (1995). Risk arbitrage in the Nikkei put warrant market, 1989–90. *Applied Mathematical Finance*, *2*, 243–271.

Swiss National Bank (2011). Introduction of a minimum swiss franc exchange rate against the euro. http://www.snb.ch/en/mmr/speeches/id/ref_20110906_pmh.

The Guardian (2015). Swiss bank's currency u-turn hurts watchmakers, skiers and traders. January 15.

Ziegel, J. (2014, published online). Coherence and elicitability. *Mathematical Finance (forthcoming)*.

Ziegler, A., & Ziemba, W. (2015). Returns from investing in S&P500 futures options. In A. Malliaris, & W. Ziemba (Eds.), *Handbook of futures markets*. Singapore, Singapore: World Scientific Publishing Co.

Ziemba, R. E., & Ziemba, W. T. (2013). *Investing in the modern age*. Singapore, Singapore: World Scientific.

Ziemba, W. T. (2003). *The stochastic programming approach to asset liability and wealth management*. Charlottesville, VA: AIMR.

Ziemba, W. T. (2007). The Russell-Yasuda Kasai InnoALM and related models for pensions, insurance companies and high net worth individuals. In S. A. Zenios, & W. T. Ziemba (Eds.), *Handbook of asset and liability management, Volume 2*, Handbooks in Finance, pp. 861–962. Amsterdam, Netherlands: North Holland.

Ziemba, W. T. (2013). The case for convex risk measures and scenario dependent correlation matrices. *Quantitative Finance Letters*, *1*(1), 47–54.

Ziemba, R. (2015). *Roubini global economics*. Private conversation. New York.

Chapter 18
Leveling the Playing Field

Jonathan B. Berk and Jules H. van Binsbergen

The development of what has become known as the Capital Asset Pricing Model (CAPM) by Jack Treynor and others was a watershed event in financial economics.[1] It marked the birth of modern asset pricing because for the first time financial economists had a formal method for estimating the expected return of a risky investment opportunity. Within a short period of time after the model was developed, researchers set about collecting stock market data to determine whether or not the model actually worked. Early results were encouraging—beta appeared to explain cross sectional differences in realized returns.[2] However, as researchers subjected the model to more powerful tests, cracks began to appear. In particular, researchers were able to group assets into portfolios using variables such as size, book-to-market, and past returns (momentum) and show that even though these portfolios displayed large cross sectional variation in realized returns, this was not mirrored in an equivalent cross sectional variation in beta. In response to these empirical

This article summarizes research originally published in Berk and van Binsbergen (2016b).

[1]The model was developed independently by Lintner (1965), Mossin (1966), Sharpe (1964), and Treynor (1961), see French (2003) for an analysis of who deserves attribution.

[2]See, Black et al. (1972) and Fama and MacBeth (1973).

J.B. Berk (✉)
Graduate School of Business, Stanford University,
Stanford, CA 94305-7298, USA

National Bureau of Economic Research
e-mail: jbberk@stanford.edu

J.H. van Binsbergen
The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA

National Bureau of Economic Research
e-mail: julesv@wharton.upenn.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_18

shortcomings, a number of extensions to the original model have been proposed. The most notable are the factor specifications proposed by Fama and French (1993) and Carhart (1997) which are motivated by theoretical developments in Ross (1976) and Merton (1973).

The new models have been subjected to the same level of empirical scrutiny as the CAPM, and for the most part they have fared better. This has led many financial economists to conclude that, to adjust for risk, the CAPM should be replaced by one of the new models. The problem with this logic is that it ignores a subtle distinction between the models. The CAPM was developed using theory alone—stock market data did not exist in electronic format when Jack wrote his original paper. The new models were developed only after observing the data. Indeed, they were actually derived with one goal in mind, to fit return data better. What that means is that they have an inherent advantage over the CAPM when they are evaluated using the data they were designed to explain. This is true even if you only test the model using data collected after the models were first proposed.

The easiest way to understand why, when you compare the performance of the new models to the CAPM, the tests are biased in favor of the new models, is to consider the following analogy. Early astronomers could not reconcile the motion of the planets with the dominant theory of the day—the Ptolemaic theory that had the Earth at the center of the universe. Rather than look for an alternative theory, these astronomers reacted to the inability of the Ptolemaic theory to predict the motion of the planets by "fixing" each observational inconsistency. Just as modern financial economists added new risk factors to the CAPM, the early astronomers added epicycles to the theory. For example, because the Ptolemaic theory did not account for the motion of the Earth, it could not explain the fact that, when viewed from the Earth, the planets sometimes move backwards. An epicycle fixes this problem by adding a circular orbit within another circular orbit. The net result was that by the time Copernicus proposed the correct theory that the Earth revolved around the Sun, the Ptolemaic theory had been fixed so many times, it *better* explained the motion of the planets than the Copernican system.[3] The lesson here is if you test a theory using the data it was designed to explain, it should not be surprising to find that the theory works. A real test of a theory is when it explains data it was not designed to explain.

Although the extensions to the CAPM better explain the cross section of stock returns, it is hard to know, using traditional tests, whether these extensions represent true progress towards a better measure of risk or simply the asset pricing equivalent of an epicycle. To determine whether any extension to the CAPM better explains risk, one needs to confront the models with facts they were not specifically designed to explain. At first glance it might appear that this approach is a lost cause. How can you test a model of risk without using stock return data? The answer is that instead of looking at stock returns, we look at what investors actually do. That is, we infer what risk model investors use by observing their investment decisions.

[3]Copernicus incorrectly assumed that the planets followed circular orbits when in fact their orbits are ellipses.

To understand the basis of our new test, it is helpful to recall how prices and returns are determined in any risk model. All models of risk assume that investors compete with each other to find attractive investment opportunities. When investors find such opportunities, they react by submitting buy or sell orders. Prices are then determined so that the market clears, that is, total demand equals total supply. As a consequence of this competition, equilibrium prices are set so that the expected return of every asset is solely a function of its risk. Consequently these buy and sell orders reveal the preferences of investors and therefore they reveal which risk model investors are using. By observing these orders we can infer whether investors price risk at all, and if so, which risk model they are using.

There are two criteria that are required to implement this idea. First, one needs a mechanism that identifies attractive investment opportunities. Second, one needs to observe investor reactions to these opportunities. We can satisfy both criteria if we implement the method using mutual fund data. Using this dataset we infer, from a set of candidate models, the model that is closest to the risk model investors are actually using. We will restrict attention to the time period after the new models were developed, that is 1996–2011. Here we follow the lead of Guerard Jr, Deng, Gillam, Markowitz, Wang, and Xu (2015) who test Bloch, Guerard Jr, Markowitz, Todd, and Xu (1993) in the 1997-2014 period.

What we find is somewhat of a triumph for economic theory. Even without the benefit of the last 50 years of data, we find that the model derived in the early 1960s, the CAPM, is the best description of investor behavior. None of the extensions that have been proposed do better. Importantly, the CAPM better explains investor behavior than no model at all, indicating that investors do price risk. Most surprisingly, the CAPM also outperforms a naive model in which investors ignore beta and simply chase any outperformance relative to the market portfolio. Investors' capital allocation decisions reveal that they adjust for risk using the CAPM beta. The poor performance of the extensions to the CAPM implies that although these extensions might better explain cross sectional variation in realized returns, they do not help explain how investors measure risk. In short, we are no closer to understanding the risk-return relation today than we were when the CAPM was originally developed more than half a century ago.

18.1 Methodology and Data

In earlier work we explain how the mutual fund market equilibrates.[4] When investors find an investment in a particular mutual fund to be attractive, they invest capital in the fund. As the fund grows, the expected return of the fund declines as the fund manager's attractive investment ideas are exhausted. The flow of capital ceases

[4]Berk and Green (2004), Berk (2005), Berk and van Binsbergen (2016b), and Berk and van Binsbergen (2016c).

when the expected return the mutual fund delivers to its investors is solely a function of the risk of the fund. That is, competition between investors drives the fund's net alpha to zero. What this implies is that the flows of capital in and out of mutual funds are the buy and sell orders mentioned in the introduction. Thus, the flow of funds reveals which investment opportunities mutual fund investors considered to be attractive.

Notice that when the market is in equilibrium, all mutual funds have a zero net alpha. Now consider what happens when new information arrives that allows investors to make a better inference about a fund's alpha. One example of new information is the fund's return itself. If the fund's return exceeds the risk adjusted return predicted by the risk model investors are using, investors will positively update their beliefs about the skill level of the fund's manager and infer that at the fund's current size, the alpha is positive. Similarly, if the fund's realized return is less than the risk adjusted return predicted by the risk model, investors will negatively update their beliefs about the skill level of the manager and infer that at the fund's current size, the alpha is negative. In short, the fund's realized return reveals attractive investment opportunities, and the subsequent flow of funds reveals investor reactions to these opportunities.

We are now ready to describe our test. Each risk model we consider uniquely determines which funds outperform and which funds underperform. We then observe the subsequent flow of funds. The model for which outperformance best drives capital flows is the model that comes closest to the model that investors are actually using to price risk. We use the mutual fund data set described in Berk and van Binsbergen (2015). Because the focus of this article is to ensure that we test all models on an equal footing, we will only conduct our test using data that was not available at the time all the models were developed. In practice that means we restrict attention to the time period from 1996–2011.[5]

We implement this idea as follows. We compute the fraction of times we observe an inflow when the fund's realized return exceeds the risk adjusted return and the fraction of times we observe an outflow when the fund's realized return is less than the risk adjusted return, as defined by the risk model. Our measure of fit is the average of these two fractions. We show in Berk and van Binsbergen (2016b) that this average can also be estimated by running a simple linear regression of the sign of flows on the sign of outperformance. The latter approach is preferable because, as we show in the same paper, the t-statistics of this regression is an accurate measure of statistical significance. In particular, if the coefficient using one risk model statistically significantly exceeds the coefficient using a second risk model, then we can say the first model is closer to the risk model investors are actually using.

[5]In a related paper, Berk and van Binsbergen (2016c), we provide an equivalent summary of the results of Berk and van Binsbergen (2016b) but for the full sample period, that is, 1977–2011.

18.2 Results

There are two practical issues that we need to confront in order to run this test. The first concerns what a flow actually is. A fund's assets under management changes for two reasons. Either the prices of the underlying stocks change or investors invest or withdraw capital. Although both mechanisms change assets under management, they are unlikely to equally affect the fund's alpha. For example, increases in fund sizes that result from inflation are unlikely to affect the alpha generating process. Similarly, the fund's alpha generating process is unlikely to be affected by changes in fund size that result from changes in the price level of the market as a whole. Consequently, in our empirical specification, we only consider capital flows into and out of funds net of what would have happened had investors not invested or withdrawn capital and had the fund manager adopted a purely passive strategy and invested in Vanguard index funds. That is, we measure the flow of funds as

$$\mathrm{SIGN}(q_{it} - q_{it-T}(1 + R_{it}^V)), \tag{18.1}$$

where q_{it} is the size of fund i at time t, and R_{it}^V is the cumulative return, over the horizon from $t - T$ to t, to investors of the collection of available Vanguard index funds that comes closest to matching the fund under consideration. Under this definition of capital flows, we are assuming that, in making their capital allocation decisions, investors take into account changes in the size of the fund that result from returns due to managerial outperformance alone. That said, all of our results are robust to replacing R_{it}^V with the fund's own return in (18.1).

The second practical issue that we need to confront is the horizon length over which to measure the effects. For most of our sample funds report their AUMs monthly, however, in the early part of the sample many funds report their AUMs only quarterly. In order not to introduce a selection bias by dropping these funds, the shortest horizon we will consider is three months. If investors react to new information immediately, then flows should immediately respond to performance and the appropriate horizon to measure the effect would be the shortest horizon possible. But in reality, there is evidence that some investors do not respond immediately. For this reason, we also consider longer horizons (up to four years). The downside of using longer horizons is that longer horizons tend to put less weight on investors who update immediately, and these investors are also the investors more likely to be marginal in setting prices.

We will consider the following models of risk. Because the market portfolio is not observable, we will test two versions of the CAPM that correspond to two different market proxies, the CRSP value weighted index of stocks and the S&P 500 index. We will also test the factor models proposed in Fama and French (1993), hereafter the FF factor specification and Carhart (1997), hereafter the FFC factor specification. In addition we will consider three "no model" benchmarks. The first uses the actual return of the fund, which corresponds to investors using no model at all. The second uses the return of the fund in excess of the risk free return. Investors

Table 18.1 *Flow of funds outperformance relationship (1996–2011)*: The table reports the average of the fraction of times we observe an inflow when the fund's realized return exceeds the risk adjusted return and the fraction of times we observe an outflow when the fund's realized return is less than the risk adjusted return. Each row corresponds to a different risk model. The first two rows report the results for the market model (CAPM) using the CRSP value-weighted index and the S&P 500 index as the market portfolio. The next three rows report the results of using as the benchmark return, three rules of thumb: (1) the fund's actual return, (2) the fund's return in excess of the risk-free rate, and (3) the fund's return in excess of the return on the market as measured by the CRSP value-weighted index. The next two rows are the FF and FFC factor specifications. The largest value in each column is shown in boldface

Model	Horizon		
	3 month	6 month	1 year
Market models (CAPM)			
CRSP value weighted	**62.74**	**62.68**	62.70
S&P 500	61.44	61.23	60.77
No model			
Return	57.94	59.48	57.45
Excess return	57.67	59.27	57.44
Return in excess of the market	61.18	61.31	60.33
Multifactor Models			
FF	62.42	62.20	**62.80**
FFC	62.57	62.35	62.71

would use this measure of risk if they were risk neutral. Finally, we will consider a model where the performance of the fund is just the fund's return minus the return of the market (as measured by the CRSP value weighted index). Although similar to the CAPM, in this model investors ignore beta. All they care about is outperformance relative to the market.

Which model best approximates the true asset pricing model? Table 18.1 reports the average of the fraction of times we observe an inflow when the fund's realized return exceeds the risk adjusted return and the fraction of times we observe an outflow when the fund's realized return is less than the risk adjusted return. If flows and outperformance are unrelated, we would expect this average to equal 50 %. The first takeaway from Table 18.1 is that none of our candidate models can be rejected,[6] implying that regardless of the risk adjustment, a flow-performance relation exists. On the other hand, none of the models perform better than 63 %. It appears that

[6]The second column of Table 18.2 reports the double-clustered (by fund and time) t-statistics under the null that flows and performance are unrelated.

a large fraction of flows remain unexplained. Investors appear to be using other criteria to make a non-trivial fraction of their investment decisions.

The CAPM with the CRSP value weighted index as the market proxy performs best at the 3- and 6-month horizon, and the FFC model performs best at the 1-year horizon. To assess whether the difference in performance between the CAPM and the other models is statistically significant, we report, in Table 18.2, the double-clustered (by fund and time) *t*-statistics. Recall that because the new models nest the CAPM, for researchers to reject those models in favor of the CAPM, they must statistically outperform the CAPM. Yet as Table 18.2 shows, no model statistically outperforms the CAPM at any horizon.

To assess the relative performance of the models, we begin by first focusing on the behavioral model that investors just react to past returns without adjusting for risk, the column marked "Ret" in the table. By looking down that column in Table 18.2, one can see that the factor models all statistically significantly outperform this model at horizons of less than two years. For example, the *t*-statistic reported in Table 18.2 that the CAPM outperforms this no model benchmark at the 3-month horizon is 4.98, indicating that we can reject the hypothesis that the behavioral model is a better approximation of the true model than the CAPM. Based on these results, we can reject the hypothesis that investors just react to past returns. The next possibility is that investors are risk neutral. In an economy with risk-neutral investors, we would find that the excess return (the difference between the fund's return and the risk free rate) best explains flows, so the performance of this model can be assessed by looking at the columns labeled "Ex. ret." Notice that all the risk models nest this model, so to conclude that a risk model better approximates the true model, the risk model must statistically outperform this model. For horizons less than 2 years, all the risk models satisfy this criterion. Finally, one might hypothesize that investors benchmark their investments relative to the market portfolio alone, that is, they do not adjust for any risk differences (beta) between their investment and the market. The performance of this model is reported in the column labeled "Ex. mkt." The CAPM statistically significantly outperforms this model at all horizons—investors' actions reveal that they use betas to allocate resources.

Next, we use our method to discriminate between the risk models. Recall that both the FF and FFC factor specifications nest the CAPM (the first factor in each specification is the market), so to conclude that either factor model better approximates the true model, it must statistically significantly outperform the CAPM. The test of this hypothesis is in the columns labeled "CAPM." Neither factor model statistically outperforms the CAPM at any horizon implying that the additional factors add no explanatory power for flows.

It is also informative to compare the tests of statistical significance across horizons. The ability to statistically discriminate between the models deteriorates as the horizon increases. This is what one would expect to observe if investors instantaneously moved capital in response to the information in realized returns. Thus, this evidence is consistent with the idea that capital does in fact move quickly to attractive investment opportunities.

Table 18.2 *Tests of statistical significance*: The first column in the table reports the average of the fraction of times we observe an inflow when the fund's realized return exceeds the risk adjusted return and the fraction of times we observe an outflow when the fund's realized return is less than the risk adjusted return. The second column provides the t-statistic of the test of whether this average is significantly different from 50 %. The rest of the columns provide the statistical significance of the pairwise test of whether the models are better approximations of the true asset pricing model. For each model in a column, the table displays the t-statistic of the test that the model in the row is a better approximation of the true asset pricing model. The rows (and columns) are ordered by the probabilities in the first column, with the best performing model on top. All t-statistics are double clustered by fund and time (see Thompson (2011))

Panel A: 3-Month horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	FFC	FF	CAPM SP500	Ex. mkt	Ret	Ex. ret
CAPM	62.74 %	21.48	0.00	0.43	0.81	4.94	6.32	4.33	4.94
FFC	62.57 %	24.38	−0.43	0.00	0.71	2.23	3.16	3.95	4.53
FF	62.42 %	23.81	−0.81	−0.71	0.00	1.92	2.76	3.87	4.45
CAPM SP500	61.43 %	16.94	−4.94	−2.23	−1.92	0.00	0.64	3.19	3.68
Excess market	61.18 %	18.39	−6.32	−3.16	−2.76	−0.64	0.00	2.88	3.32
Return	57.94 %	8.86	−4.33	−3.95	−3.87	−3.19	−2.88	0.00	1.19
Excess return	57.67 %	8.32	−4.94	−4.53	−4.45	−3.68	−3.32	−1.19	0.00

Panel B: 6-Month horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	FFC	FF	Ex mkt	CAPM SP500	Ret	Ex ret
CAPM	62.68 %	17.08	0.00	0.52	0.73	3.57	3.44	2.12	2.54
FFC	62.35 %	18.17	−0.52	0.00	0.52	1.40	1.34	1.80	2.16
FF	62.20 %	18.90	−0.73	−0.52	0.00	1.20	1.09	1.65	1.97
Excess market	61.31 %	13.00	−3.57	−1.40	−1.20	0.00	0.13	1.10	1.34
CAPM SP500	61.23 %	11.13	−3.44	−1.34	−1.09	−0.13	0.00	1.16	1.39
Return	59.48 %	7.24	−2.12	−1.80	−1.65	−1.10	−1.16	0.00	0.58
Excess return	59.27 %	6.98	−2.54	−2.16	−1.97	−1.34	−1.39	−0.58	0.00

Panel C: 1-Year horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	FF	FFC	CAPM	CAPM SP500	Ex mkt	Ret	Ex ret
FF	62.80 %	12.65	0.00	0.33	0.12	1.72	2.59	1.94	2.29
FFC	62.71 %	12.21	−0.33	0.00	0.02	1.77	2.54	1.96	2.35
CAPM	62.70 %	11.16	−0.12	−0.02	0.00	3.40	6.32	1.98	2.43
CAPM SP500	60.77 %	6.47	−1.72	−1.77	−3.40	0.00	0.48	1.35	1.57
Excess market	60.33 %	8.26	−2.59	−2.54	−6.32	−0.48	0.00	0.97	1.14
Return	57.45 %	3.54	−1.94	−1.96	−1.98	−1.35	−0.97	0.00	0.02
Excess return	57.44 %	3.46	−2.29	−2.35	−2.43	−1.57	−1.14	−0.02	0.00

18.3 Conclusion

Our empirical finding that no model outperforms the CAPM is, in some sense, startling. The model was developed at a time when the mutual fund sector was tiny. In 1962, there were just 172 equity mutual funds in existence. In the interim an entirely new sector of investing developed, so that today there are more funds then there are stocks. The other models we evaluated were all developed after the mutual fund sector started to experience explosive growth. Yet none of those models are better able to explain investor behavior. That Jack Treynor was able to predict behavior in a sector that essentially did not exist when he first developed the CAPM is a remarkable achievement in the field of economics. His subsequent application of the CAPM and beta to mutual fund performance (Treynor (1965) and Treynor and Mazuy (1966)) measurement was a great innovation in financial economics.

Yet the fact remains that the CAPM does a poor job explaining cross-sectional variation in expected returns. The profession's answer to this shortcoming has been to attempt to improve the CAPM. What our results show is that the reason these "improved" models better explain cross-sectional variation is simply because that is what they have been designed to do. We are therefore no closer at explaining what, if any, other factors determine expected returns than we were when the CAPM was first developed.

This raises a number of possibilities about the relation between risk and return. The first possibility, and the one most often considered in the existing literature, is that this finding does not invalidate the neoclassical paradigm that requires expected returns to be a function solely of risk. Instead, it merely indicates that the CAPM is not the correct model of risk, and, more importantly, a better model of risk exists.

The second possibility is that the poor performance of the CAPM is a consequence of the fact that there is no relation between risk and return. That is, that expected returns are determined by non-risk based effects. The final possibility is that risk only partially explains expected returns, and that other, non-risk based factors, also explain expected returns. The results in this paper shed new light on the relative likelihood of these possibilities.

The fact that we find that the factor models all statistically significantly outperform our "no model" benchmarks implies that the second possibility is unlikely. That leaves the question of whether the failure of the CAPM to explain the cross section of expected stock returns results because a better model of risk exists, or because factors other than risk also explain expected returns. To conclude that a better risk model exists, one has to show that the part of the variation in asset returns not explained by the CAPM can be explained by variation in risk. This is what the flow of funds data allow us to do. If variation in asset returns that is not explained by the CAPM attracts flows, as is the case for the extensions of the CAPM we tested, then one can conclude that this variation is not compensation for risk.

References

Berk, J. B. (2005). Five myths of active portfolio management. *Journal of Portfolio Management, 31*(3), 27–31.

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy, 112*(6), 1269–1295.

Berk, J. B., & van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics, 118*(1), 1–20.

Berk, J. B., & van Binsbergen, J. H. (2016a). Active managers are skilled. *Journal of Portfolio Management, 3*(2), 131–139.

Berk, J. B., & van Binsbergen, J. H. (2016b). Assessing asset pricing models using revealed preference. *Journal of Financial Economics, 119*(1), 1–23.
Berk, J. B., & van Binsbergen, J. H. (2016c). How do Investors Compute the Discount Rate? Forthcoming, Financial Analysts Journal.
Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: some empirical tests. In M. C. Jensen (Ed.) *Studies in the theory of capital markets* (pp. 79–121). New York: Praeger.
Bloch, M., Guerard Jr, J., Markowitz, H., Todd, P., & Xu, G. (1993). A comparison of some aspects of the US and Japanese equity markets. *Japan and the World Economy, 5*(1), 3–26.
Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.
Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.
Fama, E. F., & MacBeth, J. D. (1973). Risk, return and equilibrium: empirical tests. *Journal of Political Economy, 81*, 607–636.
French, C. W. (2003). The treynor capital asset pricing model. *Journal of Investment Management, 1*(2), 60–72.
Guerard Jr, J., Deng, S., Gillam, R., Markowitz, H., Wang, E., & G. Xu (2015). Data Mining Corrections Testing, presented at INFORMS, November 2015.
Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13–37.
Merton, R. C. (1973). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory, 3*, 373–413.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768–783.
Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341–360.
Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425–442.
Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics, 99*(1), 1–10.
Treynor, J. (1961). Toward a Theory of the Market Value of Risky Assets.
Treynor, J., & Mazuy, K. (1966). Can mutual funds outguess the market. *Harvard Business Review, 44*(4), 131–136.
Treynor, J. L. (1965). How to rate management of investment funds? *Harvard Business Review, 43*(1), 63–75.

Chapter 19
Against the 'Wisdom of Crowds': The Investment Performance of Contrarian Funds

Kelsey D. Wei, Russ Wermers, and Tong Yao

19.1 Introduction

In an article published in the *Financial Analysts Journal* (Treynor, 1987), Jack Treynor wrote about a series of "bean jar" experiments he conducted with students in his investments courses at the University of Southern California. In the first set of experiments, he asked students to independently estimate the number of beans contained in a full jar. While most students' individual estimates missed the actual number by a wide margin, surprisingly, the average estimates were pretty close to being correct. In the second set of experiments, he first provided students with advice on properties of the jar, such as the air space at the top of the jar, and materials of the jar. While such information supposedly could help improve the accuracy of students' estimates, the resulting average estimates, alas, had much larger errors than those from the first set of experiments. It seems his advice did nothing more than cause common errors among students!

This work draws from, and adds discussion to, our *Management Science* publication, Wei, Wermers, and Yao (2015), with permission from Informs.

K.D. Wei
Jindal School of Management, University of Texas at Dallas, Richardson, TX 75080, USA
e-mail: kelsey.wei@utdallas.edu

R. Wermers (✉)
Department of Finance, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA
e-mail: wermers@umd.edu

T. Yao
Department of Finance, Henry B. Tippie College of Business, University of Iowa, Iowa City, IA 52242, USA
e-mail: tong-yao@uiowa.edu

© Springer International Publishing Switzerland 2017

J.B. Guerard, Jr. (ed.), *Portfolio Construction, Measurement, and Efficiency*,
DOI 10.1007/978-3-319-33976-4_19

Treynor's first set of experiments was made famous by the popular book of James Surowiecki (2004) as early evidence of the "wisdom of crowds." A substantial part of Treynor's FAJ article, however, was about the second set of experiments, and their implications for potential shared errors in the stock market. He remarked that investors may be persuaded to give up their independent information and, instead, rely on certain common sources of information, such as published analyst research reports, and that this may actually do damage to market efficiency. In the FAJ article, he further contemplated a strategy to take advantage of such investor behavior, by waiting "until propagation [of the research among investors] is complete, or almost complete, and then copper it." However, he also cautioned about the considerable challenges for doing so given the difficulty of estimating the "shared error."

Treynor's notion of investors giving up their independent opinions to follow influential common advice is also known as "herding," and those who attempt to trade against herds are known as "contrarians." In this article, we examine contrarian investment behavior in the mutual fund industry, and uncover interesting empirical findings related to the characteristics and performance of contrarian funds. In particular, we find that there are mutual funds that systematically act in a contrarian fashion, as contemplated by Treynor (1987), and which are capable of delivering outperformance even after we take into account the different types of risks to which they are exposed.[1]

Prior to our study, academic researchers have focused their attention primarily on the herds—investors who follow each other in pursuing similar trades. These studies include, for example, Lakonishok, Shleifer, and Vishny (1992), Wermers (1999), Sias (2004), Dasgupta, Prat, and Verardo (2011a), and Brown, Wei, and Wermers (2014). The collective wisdom drawn from these studies is that, in less recent times (e.g., prior to the mid-1990s), mutual fund herding was relatively weak, and does not substantially distort stock prices; however, in more recent years, herding has become more prevalent, and herds tend to cause a significant price impact, followed by a return reversal.[2]

These strong results for funds that herd bring about an important question: do funds that do not herd, or that even "anti-herd," (actively invest against the crowd),exhibit different strategies and performance from their more conventional counterparts? For example, given the time-trend of increasing price impact caused by trades of herds, it is natural to wonder if sophisticated investors have emerged in recent years who choose to deviate from the crowd and take advantage of the

[1]We note that Treynor created a methodology to rate investment funds in Treynor (1965) and Treynor and Mazuy (1966), and the Fama–MacBeth (1973) regressions that we use for identifying the relation between "contrarianism" and "the abnormal returns of stocks" by contrarian funds build on this work.

[2]There is also a debate on whether herds indeed irrationally give up their own opinion and rely too much on certain influential common information sources. For example, Sias (2004) argues that herds merely infer information from each other's trades.

temporary price dislocations created by herds. Further, are contrarians "smarter" than herding funds in researching stock fundamentals? We address those questions in this article.

Anecdotal evidence suggests that some household names in the investment industry are contrarian investors. For example, in Wikipedia (which arguably represents the "wisdom of the crowd"), the entry for "contrarian investing" includes the "notable contrarian investors" Warren Buffett, Marc Faber, David Dreman, and John Neff. Successful mutual fund managers such as Peter Lynch, Bill Miller, and John Templeton have also been known to have strong contrarian elements in their investment themes. While such anecdotes are interesting, we would like to know whether a successful track-record belongs to only a rare few, or whether contrarian investing, in general, is rewarded. Also, by parsing through their positions and trades, as we do in this article, we hope to learn about the characteristics of a successful contrarian, as well as the specific sources of contrarian investor performance.

19.2 Identifying Contrarian Funds

There are various ways to define contrarian investing. For example, if we think of contrarian investing as a deep-value investing style, we can look at how funds trade on fundamental value indicators. Alternatively, we can define contrarian funds as those buying stocks whose prices have fallen dramatically. Treynor (1987) suggests influential analyst research reports as a prominent stimulus of herding by investors; thus, one can also define contrarian investing as trading against analyst recommendations. However, perhaps the most straightforward definition of contrarian funds would be those trading against herds, which is the definition that we adopt in our study. The advantage of this approach, in our belief, is that we need not assume a particular trading strategy to define contrarian funds or the specific source of common errors that they avoid (e.g., analyst reports). Instead, we can simply identify those funds that most frequently trade against herds, and let the data tell us what strategies that they tend to follow in doing so.

To construct a fund-level contrarian measure, we first classify, for a given fund, each trade into either a "herding" trade or a "contrarian" trade during a particular quarter (we infer "trades" by examining changes in quarterly portfolio holdings data available from Thomson Reuters). A herding trade is one in the same trading direction as the majority of funds (i.e., the "crowd"), while a contrarian trade is one in the opposite direction of the majority. For example, if a fund sells a stock when the majority of the funds are buying, or if a fund buys a stock when the majority of funds are selling, such a trade is contrarian; if the fund buys or sells with the crowd, that trade is a herding trade.

However, we first need a formula that identifies when a group of funds trading a stock can be considered a herd; more to the point, we need a measure of how strong or weak a herd is. Here, we rely on a stock-level measure of herding by the pioneering paper of Lakonishok, Shleifer, and Vishny (Lakonishok et al., 1992):

$$\mathrm{HM}_{i,t} = |p_{i,t} - \overline{p}_t| - E\left(|p_{i,t} - \overline{p}_t|\right), \tag{19.1}$$

where $p_{i,t}$ is the proportion of mutual funds buying stock i during quarter t, out of all funds trading that stock during quarter t. Note that $\overline{p}_t$, a proxy for the expected value of $p_{i,t}$, is the cross-sectional mean of $p_{i,t}$ over all stocks traded by all funds during quarter t. E($|p_{i,t} - \overline{p}_t|$) is an adjustment factor, which equals the expected value of $|p_{i,t} - \overline{p}_t|$ under the null of no herding.[3]

Intuitively, this measure defines herding as the tendency with which a group of funds exhibit similarity in trading activity, above what would have been expected as a result of random occurrences of same-side trading by funds in the same stock. Depending on the direction of herding, we can further define conditional buy-herding (BHM_{it}) and sell-herding (SHM_{it}) measures as follows:

$$\mathrm{BHM}_{i,t} = \mathrm{HM}_{i,t}\Big| p_{i,t} > \overline{p}_t \tag{19.2}$$

$$\mathrm{SHM}_{i,t} = \mathrm{HM}_{i,t}\Big| p_{i,t} < \overline{p}_t. \tag{19.3}$$

A positive value of *BHM* indicates that the majority of funds are buyers of the stock (hence, herding on the buy side), and a positive value of *SHM* indicates that the majority funds are sellers (hence, herding on the sell side).

Each quarter, we separately rank stocks into quintiles, based on the magnitude of *BHM* and *SHM*, and further assign negative signs to the quintile ranks of the *SHM* stocks. Thus, during a given quarter, BHM stocks are assigned ranks of 1 (least amount of buy herding) to 5 (most amount of buy herding), while SHM stocks are assigned ranks from −1 (least amount of sell herding) to −5 (most amount of sell herding). This way, we combine the buy-herding and sell-herding measures into a single variable, *HERD*, that takes on integer values from −5 to +5 (excluding 0), and that summarizes both the direction (buy or sell) and the strength of the herd.

We then measure the extent to which fund j conducts contrarian versus herding trades by computing the weighted average of the *HERD* measure across all stocks traded by that fund during a particular quarter, where the weights are proportional to the dollar values of the trades, and are denoted as ω_{ijt},

$$\mathrm{CON}_{jt} = -\sum_{i=1}^{N} \omega_{ijt}\mathrm{HERD}_{it}. \tag{19.4}$$

[3]This value is calculated assuming, under the null of no herding in stock-quarter i,t, that funds trade randomly and independently of each other. With this assumption, $p_{i,t}$ can be assumed to follow a binomial distribution with parameters (n, $\overline{p}_t$), where n = the number of funds that trade stock i during quarter t.

We term this measure the fund-level "contrarian index" or *CON*. Note that ω_{ijt} has a positive (negative) value for buy (sell) trades, whereas $HERD_{it}$ has a positive (negative) value for buy-herding (sell-herding) stocks. The value of CON is, thus, positively correlated with the (dollar) proportion of contrarian trades, and negatively correlated with the proportion of herding trades executed by a fund. In summary, a highly positive contrarian index identifies a contrarian fund, while a highly negative contrarian index identifies a herding fund.

The following example helps to illustrate the economic meaning of our definition. If almost all mutual funds are buying IBM and selling Cisco during a particular quarter, then a fund that sells IBM and buys Cisco during that quarter would be assigned a very high contrarian index. Note that this definition of contrarianism does not necessarily imply that contrarians are all alike, and form a small herd of their own. For example, some contrarians might sell IBM without buying Cisco, while others might buy Cisco without selling IBM.

19.3 Distribution and Characteristics of Contrarian/Herding Funds

Table 19.1 displays the cross-sectional distribution of the contrarian index. One eye-catching pattern is that the majority of funds have negative contrarian index (*CON*) values. The average value of the index across funds is −0.84, and even the 75th percentile is negative, at −0.33. This suggests that most funds are herds, while funds systematically pursuing strong contrarian investing constitute a relatively small group. This is not surprising, as, by definition, the majority of funds must be those that herd.

One important issue is whether the contrarian index is capable of capturing certain systematic differences in fund investment strategies, as opposed to being a mere statistical fluke. We address this issue using two different approaches. First, we ask what the distribution of *CON* would have looked like in an alternative world, where there were no intentional herding funds and no intentional contrarians. To answer this question, we randomly assign the trades observed in our data to sample funds—that is, keeping the actual trades in the data, but reshuffling the identities of which funds execute the trades. We find that, after reshuffling, the resulting contrarian indexes of individual funds exhibit a much smaller variability, relative to what are observed in the actual data. That is, we find far fewer funds heavily engaging in herding or contrarian trading in the randomized data. Thus, the distribution of the contrarian index in our data is extremely unlikely to result from random trading activities among our sample of funds, where some funds just happened to trade against the crowd frequently (by chance alone).

Second, we find that the contrarian index is quite persistent over time. Funds with high contrarian indices in one quarter tend to continue to have high contrarian indices for at least the following eight subsequent quarters. Therefore, the classi-

Table 19.1 Summary statistics

Panel A: Summary statistics on fund characteristics

	Mean	Median	Std. dev.	25th	75th
Fund_ size ($millions)	1,215	229	4,391	67	796
Total_expenses (%/year)	1.32	1.26	0.45	1.01	1.56
Turnover (%/year)	83.16	64.54	67.59	35.10	111.10
Flows (%/quarter)	1.17	−0.83	9.51	−4.02	3.92
Fund_age (years)	12.82	8.46	13.63	4.35	15.53
Raw_return (%/quarter)	2.35	2.27	4.85	−0.52	5.12
CON	−0.8374	−0.8758	0.8646	−1.3841	−0.3270

Panel B: Distributions of the stock-level herding measures and the fund-level contrarian index

	Mean	Std dev.	Min	P1	P5	P25	P50	P75	P95	P99	Max
BHM	0.0378	0.1289	−0.1371	−0.1135	−0.1060	−0.0293	0.0006	0.0824	0.2809	0.3655	0.4769
SHM	0.0376	0.1111	−0.1136	−0.1370	−0.1342	−0.0453	0.0006	0.0887	0.3193	0.3964	0.4359
CON	−0.8374	0.8646	−3.9435	−2.8706	−2.1774	−1.3841	−0.8758	−0.3270	0.6277	1.4621	2.8553

Notes. Panel A reports summary statistics for our sample of actively managed U.S. equity mutual funds from 1995 to 2012. Each quarter, we calculate the cross-sectional mean, median, standard deviation, 25th, and 75th percentile values of fund size (total net asset value), total expenses, annual turnover, quarterly flows, age, raw quarterly returns, and contrarian index. Time-series averages of these summary statistics are reported. Panel B reports detailed distributions of the stock-level herding measures and the fund-level contrarian index during the 1995–2012 period fication of funds into herding versus contrarian funds based upon our contrarian index likely reflects the purposeful pursuit of different investment strategies by some funds.

What types of funds are likely to be contrarian funds? Do they behave any differently from prior-examined funds that pursue unique strategies? In Table 19.2, we first characterize the holdings of contrarian funds. Specifically, each quarter, we sort funds into quintile portfolios based upon their contrarian indexes, then report the average characteristics of the stock holdings of each portfolio of funds. The specific holdings-based fund characteristics we report include the average size, B/M, momentum, and illiquidity quintile ranks of fund stock holdings. Table 19.2 shows that, relative to herding funds, contrarian funds tend to invest in stocks with a larger market capitalization, a higher book-to-market ratio, lower past returns, and having slightly lower liquidity. While some of these characteristics of fund holdings are consistent with various alternative definitions of “contrarianism” based upon self-designated investment styles frequently shown on fund prospectuses, we note that contrarian funds do not substantially tilt toward value stocks and low past-return stocks. Our definition of a contrarian investment strategy is, therefore, not equivalent to simple deep value investing or negative stock price feedback trading.

To further illustrate the fund characteristics associated with contrarian investing, we report fund size, expense ratio, turnover, age, past fund performance, and past flows in Table 19.2. Consistent with the idea that contrarian funds tend to be long-term investors with reduced short-term career concerns, the results indicate that contrarian funds tend to be large funds with a low portfolio turnover ratio. They also have higher risk-adjusted performance and higher Morningstar star performance ratings. Moreover, they tend to have lower performance volatility, suggesting that they are unlikely to be those with merely good recent performance—who could be expected to be able to afford to that deviate from the crowd occasionally without much risk. Consistent with their good past performance and low performance volatility, contrarian funds appear to attract much larger investor inflows than other funds.

Lastly, we contrast the contrarian index with several measures of fund strategy uniqueness examined in the literature. By construction, contrarian funds are those that deviate from the crowd, which suggest that they may be those funds that tend to deviate more from their style benchmarks. We, therefore, examine differences between our contrarian index and three prior-documented measures of fund strategy uniqueness: Industry Concentration Index (ICI; Kacperczyk, Sialm, & Zheng, 2005), Active Share (Cremers & Petajisto, 2009), and Reliance on Public Information (RPI; Kacperczyk & Seru, 2007).

Table 19.2 shows that both funds with a very low contrarian index (i.e., herding funds) and those with a very high contrarian index (i.e., contrarian funds) tend to have a greater *ICI* and *RPI*. This is not surprising, as both herding funds and contrarian funds need to take extreme positions, and, therefore, deviate from their benchmarks, even though the motivation behind their departure from the benchmarks is very different. For example, in unreported analyses, we show that, while herding funds tend to have a high *RPI* measure (as analyst recommendations

Table 19.2 Characteristics of Contrarian Funds

Panel A: Herding related characteristics

CON quintiles	CON	% Contrarian	HM	BHM	SHM
1	−1.9875	31.34	0.0810	0.0805	0.0784
2	−1.2775	37.10	0.0460	0.0449	0.0441
3	−0.8754	40.38	0.0268	0.0250	0.0256
4	−0.4442	43.81	0.0160	0.0141	0.0151
5	0.3956	52.95	0.0191	0.0168	0.0183

Panel B: Characteristics of fund holdings

CON quintiles	Size_rank	B/M_rank	MOM_rank	ILLIQ_rank	ICI	Active_share	RPI
1—Low	4.2568	2.5890	3.1263	1.3279	0.0993	0.8405	0.0985
2	4.2706	2.6076	3.1032	1.3167	0.0848	0.8236	0.0936
3	4.3338	2.6825	3.0533	1.3001	0.0784	0.8164	0.0909
4	4.3969	2.7534	2.9724	1.2858	0.0793	0.8150	0.0943
5—High	4.4763	2.8560	2.8580	1.2611	0.0947	0.8340	0.1043
High—Low	0.2196	0.2671	−0.2683	−0.0668	−0.0046	−0.0065	0.0058
	(8.04)	(11.19)	(−15.70)	(−4.55)	(−1.81)	(−1.43)	(3.38)

Panel C: Other fund characteristics

CON quintiles	TNA	Expense_ratio	Turnover	Age	Past_alpha	% Five-star	Volatility_of_ret	Past_flow	% Cash_holdings	Volatility_of_flow
1—Low	960	1.33	79.37	12.61	−0.05	8.05	1.85	1.03	4.70	2.79
2	1,081	1.32	92.27	12.69	−0.05	7.66	1.74	1.22	4.76	2.80
3	1,123	1.31	92.17	12.87	−0.06	8.17	1.66	1.53	4.78	2.85
4	1,224	1.31	83.98	12.81	−0.02	9.39	1.65	2.14	5.08	2.84
5—High	1,581	1.31	67.89	13.14	0.02	10.37	1.71	2.18	5.79	2.85
High—Low	621	−0.02	−11.48	0.53	0.07	2.32	−0.13	1.15	1.09	0.06
	(7.76)	(−1.68)	(−6.04)	(1.90)	(5.29)	(2.51)	(−6.20)	(4.37)	(8.58)	(1.07)

Notes. This table examines the characteristics of contrarian funds. Each quarter, we group funds into quintile portfolios according to their contrarian index (*CON*) and calculate the mean characteristics for each quintile, then average these means over all quarters. Panel A reports the average value of *CON*, the proportion of contrarian trades, the Lakonishok et al. (1992) herding measure (*HM*), and the buy-herding and sell-herding measures (*BHM* and *SHM*, respectively) computed among trades conducted by funds within each quintile. Panel B reports the average size (*Size_rank*); the book-to-market, momentum, and illiquidity quintile ranks (*B/M_rank*, *MOM_rank*, and *ILLIQ_rank*, respectively) of fund stock holdings; the industry concentration index (*ICI*), computed as the Herfindahl index of implied portfolio industry weights; active share (*Active_share*), measured as the share of portfolio holdings that differs from the best-fit benchmark index holdings; and reliance on public information (*RPI*), computed as the R-squared from regressing quarterly fund trades on lagged analyst recommendation changes of the underlying stocks in the past four quarters. In panel C, we report the average value of the following variables for each quintile: total net assets (in $millions) (*TNA*), expense ratio (in %) (*Expense_ratio*), turnover (in %) (*Turnover*), fund age (in years) (*Age*), Carhart (1997) four-factor alpha in the past 36 months (in % per month) (*Past_alpha*), the percentage of funds ranked as Morningstar five-star funds (*% Five-star*), standard deviation of the four-factor alpha in the past 36 months (in % per month) (*Volatility_of_ref*), prior-quarter flows (in %) (*Past_flow*), cash holdings as a percentage of total net assets (*% Cash_holdings*), and standard deviation of flows in the past 12 months (in % per month) (*Volatility_of_flow*). In panels B and C, differences in the reported variables between contrarian funds (quintile 5) and herding funds (quintile 1) and their associated t-statistics calculated with Newey–West robust standard errors are also reported are an important catalyst of herding), contrarian funds tend to trade in the opposite direction of analyst recommendations, resulting in a higher negative correlation of their trades with analyst recommendations and, thus, a higher *RPI*.[4] Lastly, the relation between the contrarian index and Active Share is also non-monotonic. In summary, we conclude that contrarianism is different from prior measures of deviation from benchmarks or fund strategy uniqueness.

19.4 Performance of Contrarian and Herding Funds

While contrarian behavior could be driven by superior private information in the context of Treynor (1987), it may also be driven by overconfidence. That is, certain fund managers might overweight their private information and underweight useful commonly observed information, due to overconfidence (Daniel, Hirshleifer, & Subrahmanyam, 1998). Under this scenario, contrarian funds would tend to underperform. Moreover, contrarian funds are likely to underperform, as well, if their departure from herds result from fund manager incentives to gamble on fund performance, as illustrated in the risk-shifting literature (Brown, Harlow, & Starks, 1996; Chevalier & Ellison, 1997; and Huang, Sialm, & Zhang, 2011).

We, therefore, compare the performance of contrarian and herding funds to gain insight into the motivation behind contrarianism. We employ three different performance measures. The first is reported net fund return, after deducting fund expenses. The second is the characteristic-adjusted abnormal return, using a method developed by Daniel, Grinblatt, Titman, and Wermers (1997). Briefly, this method calculates the abnormal returns of each stock held by a fund, then portfolio weights this abnormal return across stocks held by the fund. The abnormal return is the return of that stock, in excess of the return of an appropriate benchmark portfolio. The benchmark portfolio for a stock is the value-weighted portfolio of stocks with similar characteristics—in terms of market capitalization, book-to-market equity ratio, and price momentum—to the stock being examined. The third performance measure is the risk-adjusted fund performance based on the four-factor model of Carhart (1997). The risk-adjusted fund performance, or the "four-factor alpha," is the intercept from regressing fund returns onto four factors—the market minus T-bills factor, size, and book-to-market factors, and, additionally, a price momentum factor.

Table 19.3 shows that contrarian funds—funds ranked in the top quintile by their contrarian index—are able to generate much better performance than herding funds (those ranked in the bottom quintile). The net fund return, characteristic-adjusted return, and four-factor alpha of the contrarian funds are 2.88 %, 0.21 %, and −0.08 %, respectively, during the quarter after fund ranking. By contrast,

[4]Note that RPI is the correlation, either positive *or negative*, between fund trading and public information.

Table 19.3 Performance of Contrarian Funds

Portfolio raw returns										
CON quintiles		Quarter 1		Quarter 2		Quarter 3		Quarter 4		Cumulative
1—Low		2.77		2.19		2.05		2.06		9.31
2		2.58		2.51		2.19		2.21		9.78
3		2.72		2.44		2.33		2.33		10.06
4		2.70		2.54		2.50		2.46		10.49
5—High		2.88		2.66		2.66		2.70		11.21
High—Low		0.11		0.46		0.62		0.63		1.90
		(0.34)		(1.68)		(2.14)		(1.97)		(2.00)
	DGTW-adjusted abnormal returns					Carhart (1997) four-factor alphas of fund returns				
CON quintiles	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Cumulative	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Cumulative
1—Low	0.09	−0.17	−0.16	−0.12	−0.33	−0.44	−0.58	−0.62	−0.58	−2.12
	(0.76)	(−1.11)	(−1.17)	(−0.68)	(−1.10)	(−2.24)	(−2.74)	(−3.30)	(−2.56)	(−2.68)
2	0.02	0.08	−0.11	−0.05	−0.03	−0.50	−0.39	−0.49	−0.48	−1.80
	(0.18)	(0.53)	(−0.77)	(−0.34)	(−0.09)	(−2.49)	(−2.04)	(−2.63)	(−2.26)	(−2.36)
3	0.09	0.05	−0.01	0.01	0.16	−0.42	−0.36	−0.41	−0.35	−1.49
	(0.71)	(0.39)	(−0.05)	(0.08)	(0.56)	(−2.24)	(−2.01)	(−2.06)	(−1.96)	(−2.09)
4	0.05	0.07	0.05	0.07	0.27	−0.25	−0.19	−0.19	−0.23	−0.83
	(0.45)	(0.65)	(0.44)	(0.65)	(0.84)	(−1.32)	(−1.04)	(−1.07)	(−1.24)	(−1.19)
5—High	0.21	0.19	0.16	0.22	0.82	−0.08	−0.09	−0.02	0.07	−0.07
	(1.84)	(1.69)	(1.46)	(1.81)	(1.92)	(−0.43)	(−0.53)	(−0.09)	(0.34)	(−0.09)
High—Low	0.12	0.37	0.32	0.34	1.14	0.36	0.50	0.61	0.64	2.05
	(0.94)	(2.81)	(2.34)	(1.85)	(2.39)	(2.09)	(2.86)	(3.78)	(2.78)	(3.20)

Notes. At the end of each quarter *t*, we sort funds into quintile portfolios based on their contrarian indexes and compare their performances. We report raw returns and DGTW-characteristic-adjusted abnormal returns computed based on fund portfolio holdings, as well as Carhart four-factor alphas of reported net fund returns. Returns are reported in percentage (quarterly percentages are not annualized). We also report the performance of a zero cost portfolio that buys quintile 5 (contrarian) funds and sells quintile 1 (herding) funds; *t*-statistics calculated with Newey–West robust standard errors are in parentheses the corresponding numbers for herding funds are 2.77 %, 0.09 %, and −0.44 %, respectively. The differences between contrarian funds and herding funds in these three sets of performance measures are 0.11 %, 0.12 %, and 0.36 % per quarter.

The table also shows that performance differences remain significant for several quarters after the initial fund ranking. For example, the cumulative net return of contrarian funds is 11.21 % during the four quarters after the fund ranking, significantly higher than that of herding funds, 9.31 %.

Recall that Table 19.2 indicates that contrarian funds differ from herding funds in terms of fund size, turnover, and investor flows, as well as characteristics of fund holdings. Some of these characteristics have previously been documented to be correlated with fund performance.[5] To more robustly test whether managers of contrarian funds are truly more skilled, we perform a multivariate regression of fund performance on the contraian index, with added control variables included for these fund characteristics that may be related to fund performance. In addition, since we know that contrarian funds tend to have higher measures of strategy activeness, we also wish to control for these factors, to see whether the contrarian index has any explanatory power for performance beyond that of prior-documented measures of strategy activeness (or uniqueness). The dependent variable of this panel regression is the cumulative Carhart (1997) four-factor adjusted return for a fund over the four quarters after we measure that fund's contrarian index (and other fund characteristics).

While we do not present a table (this can be found as Table 5 in Wei, Wermers, and Yao, 2015), we find that the results from this model are consistent with the aforementioned results using our simple approach of ranking funds in Table 19.3. That is, contrarian funds consistently deliver better performance than herding funds, controlling for their differing characteristics. Specifically, a fund that buys (sells) stocks that have a buy- (sell-) herding measure that is one-quintile lower exhibits almost a 0.19 % per year higher four factor alpha during the following year.[6] Moreover, the significant return predictive power of the contrarian index remains strong after we add control variables for industry concentration, Active Share, and *RPI*, measures of strategy activeness or uniqueness that have been shown to help predict fund alphas. Thus, the success of contrarian funds is not limited to a few well-known cases, but appears to be a general phenomenon. More interestingly, the outperformance of contrarian funds suggests that contrarian managers do not appear to simply be overconfident. Given their greater past performance and inflows, and, thus, lower short-term career concerns, their contrarian trading strategies are likely motivated by their reliance on superior private information.

[5]For example, Chen, Hong, Huang, and Kubik (2004) document decreasing returns-to-scale among mutual funds.

[6]Recall that buying stocks (selling stocks) with a lower buy-herding (sell-herding) measure means that the fund tends to trade against the crowd; i.e., the fund is more contrarian in its trading behavior.

19.5 What Does It Take to Be A Successful Contrarian? Parsing Through Fund Trades

What is the source of outperformance by contrarian funds? Prior studies (e.g., Dasgupta et al., 2011a and Dasgupta, Prat, & Verardo, 2011b; Brown et al., 2014) show that fund herding results in a significant short-term price impact that tends to reverse in the long-run. Is it that contrarian funds simply profit from the temporary price pressure effect of herding? If so, it seems that many investors could potentially mimic their success by simply taking the opposite position of mutual fund herds. On the other hand, if contrarian funds profit from their superior information, they should outperform, regardless of whether they trade with or against herds.

To answer this question, we parse through fund trades to examine what types of trades contribute to contrarian fund outperformance. We break down all fund trades into 40 (5x2x4) groups. First, we classify funds, by their contrarian index, into quintiles. Then, within each contrarian index quintile, we group fund trades, by direction, into buy and sell trades. Finally, within each fund quintile rank and trade direction category, we further break trades into four types, depending on the contrarian/herding nature of the trades. Type 1 consists of contrarian trades of strongly herded stocks, Type 2 for contrarian trades of weakly herded stocks, Type 3 for herding trades of strongly herded stocks, and Type 4 for herding trades of weakly herded stocks. A stock is considered a "strong herding stock" if either its buy-herding measure (BHM) or sell herding (SHM) measure is ranked in the top two BHM or SHM quintiles, respectively, among all stocks during the same quarter; otherwise the stock is considered a "weak herding stock. We then report the quarter-by-quarter performance of the resulting 40 trade portfolios (5 fund quintiles, 2 trade directions, and 4 trade types) during the following four quarters.

Table 19.4 displays the quarter-by-quarter, as well as cumulative abnormal returns (characteristic-adjusted returns described earlier; for robustness, the four-factor alphas of the return difference between contrarian and herding funds is also presented) during the 4 quarters after trading, of the 40 different types of trades. Let us focus on the buy trades first. Contrarian funds outperform herding funds on Type-1 buys, i.e., contrarian buys of stocks strongly sold by herds. Consistent with the documented short-term price impact of institutional herding, contrarian fund Type-1 buys initially do not outperform during the first quarter, but significantly outperform starting from the second quarter.

Interestingly, contrarian funds also outperform (relative to the same types of trades by herding funds) in the other three types of buy trades—contrarian buys of weakly herded stocks (Type 2), and herding buys on strongly and weakly herded stocks (Type 3 and 4). For example Contrarians also outperform in their contrarian buys of weak herding stocks (i.e., type-2 trades), where the profit from riding on the reversals of the price-pressure effect is likely small. More interestingly, contrarian fund buy trades outperform those of herding funds, even when they trade with herds (Types 3 and 4 trades). Specifically, while herding funds experience significantly negative returns in their herding trades of strong-herding stocks during quarters

Table 19.4 Trade-Based Performance of Contrarian Funds

		Buy trades					Sell trades				
	CON	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Cumulative	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Cumulative
Type 1 trades	1—Low	0.06	−0.30	0.01	0.31	0.09	0.26	−0.24	−0.65	−0.50	−1.00
Contrarian trades of strong herding stocks		(0.22)	(−1.08)	(0.04)	(1.04)	(0.17)	(0.83)	(−0.77)	(−1.96)	(−1.29)	(−1.62)
	2	−0.18	0.11	0.12	0.44	0.74	0.20	−0.15	−0.24	−0.63	−0.82
		(−0.66)	(0.38)	(0.45)	(1.48)	(1.71)	(0.66)	(−0.47)	(−0.71)	(−1.49)	(−1.48)
	3	0.04	0.24	0.49	0.75	1.58	0.48	0.05	−0.44	−0.33	−0.24
		(0.16)	(1.02)	(1.77)	(2.77)	(3.07)	(1.66)	(0.15)	(−1.66)	(−1.00)	(−0.59)
	4	−0.10	−0.02	0.36	0.27	0.56	0.49	0.00	−0.38	−0.10	0.04
		(−0.46)	(−0.09)	(1.50)	(1.21)	(1.07)	(1.59)	(0.00)	(−1.53)	(−0.33)	(0.08)
	5—High	−0.11	0.13	0.55	0.75	1.57	0.64	−0.04	−0.05	−0.41	0.13
		(−0.55)	(0.61)	(2.17)	(3.07)	(2.18)	(2.39)	(−0.12)	(−0.18)	(−1.51)	(0.27)
DGTW	High—Low	−0.18	0.43	0.54	0.44	1.48	0.38	0.21	0.60	0.09	1.13
		(−0.57)	(1.75)	(1.78)	(1.69)	(1.73)	(1.55)	(0.92)	(2.57)	(0.33)	(2.00)
Carhart	High—Low	−0.12	0.70	0.84	0.69	1.96	0.60	0.42	0.69	0.00	1.61
		(−0.29)	(2.38)	(2.22)	(1.96)	(1.92)	(1.91)	(1.36)	(2.37)	(0.00)	(2.21)

Type 2 trades	1—Low	−0.17	−0.21	−0.06	0.38	−0.17	0.19	0.02	0.11	0.16	0.60
Contrarian trades of weak herding stocks		(−0.67)	(−0.87)	(−0.26)	(1.54)	(−0.44)	(0.81)	(0.09)	(0.57)	(0.65)	(1.15)
	2	−0.21	−0.12	0.01	0.14	−0.09	0.31	0.15	0.13	0.13	0.85
		(−0.95)	(−0.51)	(0.06)	(0.67)	(−0.23)	(1.41)	(0.55)	(0.58)	(0.60)	(1.28)
	3	−0.23	−0.12	−0.10	0.53	0.03	0.14	0.12	0.03	−0.11	0.27
		(−1.14)	(−0.63)	(−0.45)	(2.77)	(0.07)	(0.69)	(0.55)	(0.15)	(−0.47)	(0.59)
	4	−0.22	0.08	0.02	0.50	0.41	0.19	0.10	0.03	0.19	0.62
		(−1.11)	(0.40)	(0.08)	(2.33)	(0.89)	(0.99)	(0.52)	(0.19)	(0.90)	(1.39)
	5—High	0.23	0.29	0.27	0.55	1.31	0.20	0.28	0.11	0.12	0.78
		(1.38)	(1.36)	(1.22)	(2.37)	(2.23)	(1.06)	(1.31)	(0.60)	(0.53)	(1.55)
DGTW	High—Low	0.40	0.50	0.33	0.17	1.47	0.01	0.26	0.00	−0.04	0.18
		(1.89)	(2.58)	(1.32)	(0.69)	(2.43)	(0.04)	(1.02)	(0.02)	(−0.22)	(0.44)
Carhart	High—Low	0.78	0.74	0.65	0.57	2.74	0.47	0.37	0.27	0.24	1.43
		(3.01)	(2.68)	(2.34)	(1.69)	(3.53)	(1.81)	(1.29)	(1.09)	(0.80)	(2.20)

(continued)

Table 19.4 (continued)

		Buy trades					Sell trades				
	CON	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Cumulative	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Cumulative
Type 3 trades	1—Low	0.26	−0.42	−0.63	−0.67	−1.46	−0.60	−0.10	0.46	0.56	0.38
Herding trades of strong herding stocks		(0.80)	(−1.17)	(−2.18)	(−1.81)	(−2.42)	(−2.70)	(−0.42)	(1.65)	(2.10)	(0.87)
	2	0.21	−0.03	−0.44	−0.51	−0.76	−0.19	−0.05	0.24	0.29	0.38
		(0.73)	(−0.07)	(−1.58)	(−1.50)	(−1.38)	(−0.83)	(−0.27)	(0.88)	(1.38)	(0.99)
	3	0.21	0.03	−0.26	−0.40	−0.36	−0.10	−0.14	0.42	0.44	0.63
		(0.75)	(0.10)	(−1.06)	(−1.30)	(−0.81)	(−0.51)	(−0.67)	(2.00)	(1.91)	(1.31)
	4	0.23	0.01	−0.22	−0.19	−0.16	−0.13	0.06	0.30	0.60	0.87
		(0.79)	(0.03)	(−0.94)	(−0.66)	(−0.40)	(−0.70)	(0.40)	(1.41)	(2.95)	(2.25)
	5—High	0.46	0.17	−0.08	0.10	0.72	−0.01	0.22	0.47	0.49	1.34
		(1.94)	(0.83)	(−0.35)	(0.37)	(1.42)	(−0.03)	(1.20)	(2.15)	(2.30)	(2.74)
DGTW	High—Low	0.21	0.59	0.55	0.77	2.18	0.60	0.32	0.01	−0.08	0.95
		(0.90)	(2.32)	(2.17)	(2.55)	(3.17)	(2.46)	(1.34)	(0.03)	(−0.27)	(1.57)
Carhart	High—Low	0.51	0.88	1.05	1.01	3.44	0.63	0.61	0.28	0.28	1.96
		(1.54)	(2.85)	(3.00)	(2.78)	(4.07)	(2.33)	(2.27)	(0.96)	(0.75)	(2.64)

Type 4 trades	1—Low	0.23	−0.20	−0.14	−0.34	−0.32	−0.25	−0.15	0.11	0.40	0.12
Herding trades of weak herding stocks		(1.02)	(−0.72)	(−0.63)	(−1.52)	(−0.67)	(−1.11)	(−0.60)	(0.52)	(1.76)	(0.25)
	2	0.24	0.21	−0.09	−0.10	0.33	−0.05	0.14	−0.20	0.35	0.28
		(1.16)	(0.89)	(−0.43)	(−0.52)	(0.67)	(−0.26)	(0.86)	(−0.99)	(1.74)	(0.65)
	3	0.20	0.20	0.04	0.03	0.58	0.03	0.20	−0.01	0.37	0.75
		(1.05)	(0.82)	(0.20)	(0.14)	(1.26)	(0.15)	(0.99)	(−0.06)	(1.70)	(1.29)
	4	0.27	0.15	−0.04	0.26	0.71	−0.04	0.03	−0.04	0.26	0.22
		(1.43)	(0.76)	(−0.18)	(1.37)	(1.48)	(−0.22)	(0.17)	(−0.21)	(1.09)	(0.41)
	5—High	0.46	0.09	0.27	0.34	1.25	−0.08	0.29	0.14	0.54	0.97
		(2.27)	(0.47)	(1.30)	(1.55)	(2.19)	(−0.37)	(1.45)	(0.70)	(2.64)	(1.83)
DGTW	High—Low	0.23	0.29	0.42	0.68	1.57	0.17	0.44	0.03	0.15	0.86
		(0.95)	(1.20)	(1.89)	(2.33)	(2.85)	(0.70)	(1.74)	(0.13)	(0.68)	(1.63)
Carhart	High—Low	0.52	0.67	0.91	0.92	2.89	0.63	0.55	0.15	0.38	1.69
		(1.70)	(2.26)	(3.25)	(2.57)	(3.66)	(2.69)	(1.93)	(0.59)	(1.22)	(2.62)

Notes. Each quarter we sort funds into quintile portfolios based on their contrarian indexes. Within each fund, we break down fund trades into four types: (1) contrarian trades of strong herding stocks, (2) contrarian trades of weak herding stocks, (3) herding trades of strong herding stocks, (4) herding trades of weak herding stocks. We then measure the average quarterly and four-quarter cumulative DGTW-adjusted abnormal returns of each type of trade within each quintile portfolio. Returns are reported in %/quarter. We also report the performance difference in DGTW-adjusted abnormal returns and Carhart (1997) four-factor alphas between quintile 5 (contrarian) funds and quintile 1 (herding) funds; t-statistics calculated with Newey–West robust standard errors are in parentheses

$t + 3$ and $t + 4$ (when the initial price pressure of fund herding reverses), contrarian funds generate zero abnormal returns on those trades. Therefore, contrarians trade on the same side as the crowd for certain stock when their own private information conforms to that of the crowd. In this case, herding is associated with a permanent price impact.

These findings suggest that, although, by construction, contrarian funds are more likely to trade away from the crowd; they do not just mechanically trade against the crowd. In fact, contrarians often end up trading with herds, as a significant portion of their trades are in the same direction as herds (Table 19.2). Therefore, the success of contrarian funds is not merely due to taking advantage of the price-pressure effect of herding (i.e., their contrarian trades). They are likely to have profited from their own source of information, even though such information may or may not conform to that of the crowd.

Next, we turn to the performance of the sell trades. The exhibit shows that there exist very small performance differences between contrarian funds and herding funds among Type-2 and Type-3 sell trades, although stocks sold by contrarian funds actually earn higher returns than stocks sold by herding funds in their Type-1 and Type-4 trades. In addition, unlike buy-trades, returns to sell-trades do not follow a particular time pattern. This result is consistent with the findings of previous studies (e.g., Chen, Jegadeesh, & Wermers, 2000; Wermers et al., 2012) that stocks sold by skilled funds tend to have *higher* returns than stocks sold by funds deemed unskilled. Since mutual funds generally do not short-sell stocks, the stocks they sell to finance purchases of other attractive stocks must come from their existing holdings. While the stocks contrarian funds sell may be expected to underperform those they buy given their superior overall performance, such stocks may not necessarily underperform those held or sold by herding funds, if the latter funds are less skillful in selecting stocks to begin with. In addition, sell trades of contrarian funds may be driven by liquidity needs (to meet investor flows) as well as to fund even more attractive stock purchases.

Overall, the trade-based analysis reveals that contrarian managers do not simply benefit from, mechanically, the price-pressure caused by fund herding. Rather, they appear to possess superior private information, as they trade independently and may end up trading with or against the crowd, depending on whether their information conforms to that of other funds. Such private information, rather than overconfidence or gambling incentives, is likely to be the source of their contrarian trading behavior and consequently their outperformance.

19.6 Extracting Stock Selection Information From Contrarian Fund Holdings

Following the observation that contrarian funds may possess superior private information, we further extract such information and aggregate it into a stock selection signal. To do so, we adopt an approach developed by Wermers et al.

(2012) to construct a stock-level contrarian score from fund holdings and the fund contrarian index. This contrarian score measures the relative degree to which a stock is held by contrarian funds versus herding funds. Intuitively, if a stock is held heavily by contrarian funds and held lightly by herding funds, this score is high. But, if a stock is held equally by contrarian funds and herding funds, the score is neutral. Intuitively, since contrarian funds possess superior investment skills, their investment choices as extracted from their portfolio holdings can be used to locate stocks with superior future returns.

Specifically, we construct a stock level contrarian score by adopting the fund level contrarian index as the fund skill proxy in Wermers et al. (2012). In our setting, the generalized inverse approach in Wermers et al. (2012) leads to the following stock-level contrarian score:

$$\alpha_{\text{CON}} = \left(\mathbf{V}'\mathbf{D}^{+}\mathbf{V}\right)\mathbf{X}'\mathbf{CON}, \tag{19.5}$$

where **CON** is the M X 1 vector consisting of elements CON_{jt} (the fund-j contrarian index score at the end of quarter t), **X** is the M X N matrix of fund portfolio weights, x_{ijt}, **V** is the first K eigenvectors of $\mathbf{X}'\mathbf{X}$, corresponding to the K largest eigenvalues. $\mathbf{D}^{+}$ is a M X M diagonal matrix whose first K diagonal elements are the inverse of the largest K eigenvalues of $\mathbf{X}'\mathbf{X}$, with the remaining M-K diagonal elements being zeros. Following Wermers et al. (2012), K is set to M/2. The higher a stock's contrarian score, α_{CON}, the more heavily the stock is held by contrarian funds, as oppossed to herding funds. If contrarian funds possess superior investment skills, we would expect stocks with a higher α_{CON} score to earn higher abnormal returns in the future.

Before we examine this prediction, we compare the stock-level contrarian score with various quantitative stock selection factors, in order to understand whether contrarian fund investment strategies are systematically related to certain stock characteristics that also help to predict stock returns. Table 19.5 shows that stocks with higher contrarian scores have stronger value-oriented characteristics, fewer investment and financing activities, higher operating efficiency, more intangible investments, and greater illiquidity. Further, they have lower earnings momentum, higher uncertainty, and lower profitability. By and large, these results are consistent with the view that contrarian funds prefer value stocks and shy away from glamorous, profitable, or liquid stocks.

Lastly, we conduct regression analyses to examine how much price-pressure, public valuation signals, or private information each contribute to the superior performance of stocks preferred by contrarian funds. Specifically, we perform Fama–MacBeth regressions of DGTW-adjusted abnormal returns of stocks, during each of the four quarters after we measure the contrarian score, on their contrarian score, controlling for the price-pressure effect associated with herding, and the various valuation signals that are correlated with the contrarian score. We show, in Table 19.6, that the contrarian score significantly predicts stock returns during the subsequent four quarters after signal construction. The return-predictive power of the contrarian score is robust to controlling for the price impact of herding funds,

Table 19.5 Contrarian Score, Herding Intensity, and Quantitative Stock Characteristics

	HERD (Q 0)	HERD (Q−1)	HERD (Q−2)	HERD (Q−3)	GIV	VAL	INVFN	EQAL	EFF	INTAG	EMOM	PROF	UNCT	ILLIQ
D1—Low	0.58	0.63	0.59	0.47	−0.08	45.71	42.32	49.45	48.79	49.76	55.62	59.92	57.44	27.27
D2	0.54	0.60	0.52	0.45	0.02	47.00	43.70	48.77	49.66	50.42	52.88	54.80	54.23	35.48
D3	0.51	0.49	0.49	0.40	0.06	47.80	45.08	48.91	50.11	50.53	50.72	51.38	51.92	41.35
D4	0.29	0.42	0.41	0.38	0.04	48.28	46.51	49.18	50.32	50.72	49.19	48.75	49.68	46.47
D5	0.21	0.23	0.30	0.35	0.06	48.65	48.22	49.34	50.58	51.04	48.53	46.88	48.02	51.23
D6	0.01	0.15	0.27	0.26	0.07	49.59	50.88	49.08	49.93	50.41	47.78	44.80	46.59	58.08
D7	−0.13	0.14	0.26	0.29	0.03	53.06	54.93	49.34	48.72	47.79	47.77	43.82	45.65	67.03
D8	−0.20	0.11	0.29	0.21	0.07	49.02	49.80	49.39	50.36	51.85	46.45	45.15	45.65	54.26
D9	−0.12	0.02	0.07	0.16	0.08	49.03	48.16	49.91	50.01	52.35	47.27	49.62	49.02	42.18
D10—High	−0.32	−0.24	−0.15	−0.07	0.19	50.70	49.02	50.72	49.08	51.80	48.58	59.65	53.95	28.55
High—Low	−0.90	−0.87	−0.74	−0.55	0.27	4.98	6.71	1.27	0.29	2.04	−7.04	−0.27	−3.48	1.27
	(−13.29)	(−16.88)	(−17.85)	(−13.78)	(2.77)	(5.51)	(8.34)	(1.67)	(0.76)	(3.43)	(−11.28)	(−0.48)	(−5.51)	(2.67)

Notes. In each quarter t (denoted as Q0), we sort stocks into deciles based on their contrarian score α_{CON}. For each decile we calculate the average herding index (*HERD*) for the four quarters from quarter $t-3$ (denoted as Q–3) to quarter t (denoted as Q0), as well as 10 categorical stock characteristic measures. *HERD* is a stock's signed herding intensity measure based on its quintile ranks of buy-herd and sell-herd measures. *GIV* is the generalized inverse alpha of Wermers, Yao, and Zhao (2012). *VAL* is a value investment measure. *INVFIN* is a measure of investment and financing activities. *EQAL* is a measure of earnings quality. *EFF* is a measure of operating efficiency. *INTAG* is a measure of intangible investment. *EMOM* is a measure of earnings momentum. *PROF* is a measure of profitability. *UNCT* is a measure of uncertainty. *ILLIQ* is a measure of illiquidity. These measures are constructed by averaging over the percentile ranks of the underlying variables, the details of which are provided in §A.1 of the appendix (where 100 % means the highest rank). The underlying variables of these categorical measures are signed so that a higher value of each categorical measure is associated with higher subsequent stock returns as suggested in the existing literature. We also report differences in herding intensity and stock characteristics between the top and bottom stock deciles along with their corresponding t-statistics (in parentheses) calculated with Newey–West robust standard errors

Table 19.6 Contrarian Score and Stock Returns: Controlling for Herding and Return-Predictive Stock Characteristics

Explanatory variables	(1)	(2)	(3)	(4)
α_{CON}	0.0090	0.0072	0.0063	0.0049
	(8.46)	(7.22)	(6.05)	(3.36)
HERD (Q 0)		−0.0439		−0.0481
		(−2.42)		(−2.22)
HERD (Q − 1)		−0.0721		−0.0794
		(−4.35)		(−4.34)
HERD (Q − 2)		−0.0639		−0.0680
		(−4.15)		(−3.91)
HERD (Q − 3)		−0.0399		−0.0557
		(−2.53)		(−2.96)
GIV			3.3926	3.6866
			(4.24)	(3.74)
VAL			−0.0055	−0.0062
			(−1.55)	(−1.81)
INVFN			−0.0026	−0.0041
			(−0.77)	(−1.21)
EQAL			0.0039	0.0038
			(3.46)	(3.37)
EFF			0.0346	0.0343
			(9.67)	(9.67)
INTAG			0.0219	0.0216
			(5.31)	(5.28)
EMOM			0.0005	0.0023
			(0.25)	(1.15)
PROF			−0.0159	−0.0171
			(−2.80)	(−3.06)
UNCT			0.0095	0.0097
			(2.68)	(2.70)
ILLIQ			0.0135	0.0141
			(3.54)	(3.76)
R-squared	0.0004	0.0024	0.0243	0.0255

Notes. This table reports coefficients from quarterly Fama–MacBeth regressions of individual stocks' DGTW-characteristic-adjusted stock returns in each of the four quarters after portfolio formation (quarter +1, quarter +4) on α_{CON}. Coefficients reported in the table, following the "JT4" overlapping portfolio approach, are those averaged over four different regressions with stock returns (the dependent variable) in the same quarter, but the explanatory variables measured over each of the past four quarters. The main explanatory variable is cross-sectional percentile rank of the contrarian score for individual stocks, α_{CON}. The control variables include the adjusted herding intensity measure *HERD* in the most recent four quarters (quarter −3, quarter 0), the generalized alpha from Wermers et al. (2012), and nine categorical stock characteristics measured at the portfolio formation quarter (quarter 0). To avoid a significant reduction of sample size, missing quantitative stock characteristics are replaced by simulated values using a multiple imputation procedure and time-series *t*-statistics reported in parentheses are adjusted to account for such simulated regressors; *R*-squared is the average adjusted *R*-squared of the Fama–MacBeth regressions as well as well-known quantitative stock selection factors. These stock-level results further confirm our conjecture that the stock selection information possessed by contrarian funds is private, and goes beyond the mere exploitation of price-pressure caused by herds or publicly available quantitative signals.

19.7 Conclusions

A short article by Treynor (1987) offers insights on potential mispricing caused by investors' herding behavior, and muses on strategies to take advantage of such mispricing. The findings of our recent study echo his insights. We identify contrarian and herding mutual funds and examine their characteristics, performance, and trades. We find that contrarian funds outperform herding funds by a significant margin. The success of the contrarian funds depends in part on their contrarian trades against herds. However, it appears that contrarian funds also possess private stock selection information. Thus, merely mimicking their contrarian trades will not make one as successful.

References

Brown, K. C., Harlow, W. V., & Starks, L. T. (1996). Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry. *Journal of Finance, 51*, 85–110.

Brown, N. C., Wei, K. D., & Wermers, R. (2014). Analyst recommendations, mutual fund herding, and overreaction in stock prices. *Management Science, 60*, 1–20.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.

Chen, J., Hong, H., Huang, M., & Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review, 94*, 1276–1302.

Chen, H. L., Jegadeesh, N., & Wermers, R. (2000). An examination of the stockholdings and trades of fund managers. *Journal of Financial and Quantitative Analysis, 35*, 43–68.

Chevalier, J., & Ellison, G. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy, 105*, 1167–1200.

Cremers, M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies, 22*, 3329–3365.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance, 52*, 1035–1058.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and over-reactions. *Journal of Finance, 53*, 1839–1886.

Dasgupta, A., Prat, A., & Verardo, M. (2011a). The price impact of institutional herding. *Review of Financial Studies, 24*, 892–925.

Dasgupta, A., Prat, A., & Verardo, M. (2011b). Institutional trade persistence and long-term equity returns. *Journal of Finance, 66*, 635–663.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*, 607–636.

Huang, J., Sialm, C., & Zhang, H. (2011). Risk-shifting and mutual fund performance. *Review of Financial Studies, 24*, 2575–2616.

Kacperczyk, M., & Seru, A. (2007). Fund manager use of public information: New evidence on managerial skills. *Journal of Finance, 62*, 485–528.

Kacperczyk, M., Sialm, C., & Zheng, L. (2005). On the industry concentration of actively managed equity mutual funds. *Journal of Finance, 60*, 1983–2011.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics, 32*, 23–43.

Sias, R. W. (2004). Institutional herding. *Review of Financial Studies, 17*, 165–206.

Surowiecki, J. (2004). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations. Doubleday.

Treynor, J. (1965). How to rate the management of investment funds. *Harvard Business Review, 43*, 63–75.

Treynor, J. (1987). Market Efficiency and Bean of Jar Experiment *Financial Analyst, 43*, 50–53.

Treynor, J., & Mazuy, K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131–136.

Wei, K. D., Wermers, R., & Yao, T. (2015). Uncommon value: The characteristics and investment performance of contrarian funds. *Management Science, 61*, 2394–2414.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *Journal of Finance, 54*, 581–622.

Wermers, R., Yao, T., & Zhao, J. (2012). Forecasting stock returns through an efficient aggregation of mutual fund holdings. *Review of Financial Studies, 25*, 3490–3529.