Advances in Japanese Business and Economics 4

Tsutomu Watanabe
Iichiro Uesugi
Arito Ono *Editors*

The Economics of Interfirm Networks

Springer

Advances in Japanese Business and Economics

Volume 4

Advances in Japanese Business and Economics showcases the research of Japanese scholars. Published in English, the series highlights for a global readership the unique perspectives of Japan's most distinguished and emerging scholars of business and economics. It covers research of either theoretical or empirical nature, in both authored and edited volumes, regardless of the subdiscipline or geographical coverage, including, but not limited to, such topics as macroeconomics, microeconomics, industrial relations, innovation, regional development, entrepreneurship, international trade, globalization, financial markets, technology management, and business strategy. At the same time, as a series of volumes written by Japanese scholars, it includes research on the issues of the Japanese economy, industry, management practice and policy, such as the economic policies and business innovations before and after the Japanese "bubble" burst in the 1990s.

Overseen by a panel of renowned scholars led by Editor-in-Chief Professor Ryuzo Sato, the series endeavors to overcome a historical deficit in the dissemination of Japanese economic theory, research methodology, and analysis. The volumes in the series contribute not only to a deeper understanding of Japanese business and economics but to revealing underlying universal principles.

More information about this series at http://www.springer.com/series/11682

Tsutomu Watanabe • Iichiro Uesugi • Arito Ono
Editors

The Economics of Interfirm Networks

Springer

Editors
Tsutomu Watanabe
Graduate School of Economics
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo
Japan

Iichiro Uesugi
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitachi
Tokyo
Japan

Arito Ono
Faculty of Commerce
Chuo University
742-1 Higashinakano, Hachioji-shi
Tokyo
Japan

ISSN 2197-8859 ISSN 2197-8867 (electronic)
Advances in Japanese Business and Economics
ISBN 978-4-431-55389-2 ISBN 978-4-431-55390-8 (eBook)
DOI 10.1007/978-4-431-55390-8

Library of Congress Control Number: 2015937372

Springer Tokyo Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgements

This volume is a product of the research project titled "Designing industrial and financial networks to achieve sustainable economic growth" funded under Japan's Ministry of Education, Culture, Sports, Science and Technology program called "Promoting social science research aimed at solutions of near-future problems" in place during 2008–2013. We thank the Ministry of Education, Culture, Sports, Science and Technology for its financial support.

Our research project would not have been successfully implemented without the collaboration between Hitotsubashi University where the project leaders—Watanabe (2008–2011, now at the University of Tokyo) and Uesugi (2011–2013)—worked/works and Teikoku Databank Ltd. (TDB), which provided us with a unique and massive database. Particularly, we thank Taro Aihara, Takuro Kito, Goro Komatsuzaki, Masahiro Miyatani, and Takashi Suzuki for their understanding on the relevance of our research project and their technical support on the database. The collaborations during the implementation of two firm-level surveys that focused on interfirm networks and a survey on the relationships between firms and finance companies are also gratefully acknowledged. Of these three surveys, the 2009 survey on interfirm networks was jointly implemented with the Small and Medium Enterprise Agency of Japan's Ministry of Economy, Trade and Industry and the Ota City Office, and we thank Masashi Aida, Hiroki Aoki, Mitsuaki Hoshino, Yoshiaki Ishii, Kei Nara, and Masato Yamada for their survey endeavors. Toshiaki Funakubo, Masaru Inaba, Yukihiro Miyazawa, Koichi Mukoyama, Yudai Nakamura, Chihiro Shimizu, and Moriya Takemoto provided useful feedback on the survey questionnaires. Last but not least, we thank the over 3000 firms that kindly responded to the three surveys.

Earlier drafts of some of the chapters in this volume were presented at the international workshop held on November 29–30, 2012, in Tokyo, jointly hosted by Hitotsubashi University, TDB, and the Research Institute of Economy, Trade and Industry (RIETI).[1] It is our great pleasure that the two keynote speakers at this international conference—Gilles Duranton and Hans Degryse—also contributed to

[1] The workshop program and presentation materials are available at: http://www.rieti.go.jp/en/events/12112901/info.html.

this volume. We thank Vasco Carvalho, Junichi Nishimura, and Yoshiaki Ogura for presenting their research and Masahisa Fujita, Yoshiaki Fujiwara, Kyoji Fukao, Masahiro Hori, Yoshitsugu Kanemoto, Keiichiro Kobayashi, Hitoshi Mitsuhashi, Masayuki Morikawa, Atsushi Nakajima, Koji Nakamura, Makoto Nirei, Mitsukuni Nishida, Yukako Ono, Michio Suzuki, Shumpeter Tamada, Yukihiro Yasuda, and other participants for their useful comments. All those mentioned provided helpful insights that significantly improved the quality of the volume.

Finally, the Hitotsubashi University-TDB-RIETI international conference was preceded by two workshops and a number of small research meetings organized by the project itself and by the RIETI study group on corporate finance and interfirm networks, all of which formed a foundation for the success of the project. We thank Sadao Nagaoka and Ryo Nakajima for insightful comments at both workshops and Haruhiko Ando, Hikaru Fukanuma, Masazumi Hattori, Hideaki Hirata, Tomohiko Inui, Takaya Kimiwada, Yusuke Naito, Tadanobu Nemoto, Shuhei Nishitateno, Keiichiro Oda, Shunsuke Okubo, Takaaki Ohnishi, Koji Sakai, Hideto Sato, Hideki Takayasu, Misako Takayasu, Suguru Tamura, Daisuke Tsuruta, Taisuke Uchino, Hayafumi Watanabe, Wako Watanabe, Peng Xu, and Nobuyoshi Yamori for their active participation and constructive comments on various occasions.

Needless to say, any remaining errors are our responsibility. The views expressed in this volume are those of the authors and do not necessarily reflect those of the institutions with which they are affiliated.

Contents

Part III Bank-Firm Relationships and Firm Dynamics

About the Editors

Tsutomu Watanabe has been a professor of economics at the Graduate School of Economics, The University of Tokyo, since October 2011. He was a professor at Hitotsubashi University (1999–2011) and a senior economist at the Bank of Japan (1982–1999). He has held visiting positions at various universities, including Kyoto University, Bocconi University, and Columbia University. He received his Ph.D. in economics from Harvard University in 1992 and did his undergraduate work at The University of Tokyo. Watanabe's main research area is monetary policy and inflation dynamics. He is known for his series of papers on monetary policy when nominal interest rates are bounded at zero; in particular, his paper on the optimal monetary policy at the zero lower bound has been widely recognized as the first paper to provide a simple description of the liquidity trap and characterize the optimal policy response to it in a setting of dynamic stochastic general equilibrium. He is an author of many books and more than 40 academic journal articles on monetary policy, inflation dynamics, and international finance.

Iichiro Uesugi is a professor at the Institute of Economic Research of Hitotsubashi University, Japan. He received his M.A. and Ph.D. in economics in 2000 from the University of California, San Diego. Prior to becoming a member of the faculty of Hitotsubashi University in 2011, he joined the Ministry of Economy, Trade and Industry (METI) in 1993 and worked in the Economic Policy Unit of the Minister's Secretariat, Macroeconomic Affairs Division; the Planning Division of the Coal Department in the Agency of Natural Resources and Energy; the Research Division of the Small and Medium Enterprises Agency; and the Research Institute of Economy, Trade and Industry. His research interests focus on firm dynamics, financial intermediation, and the Japanese economy. His research has been published in a number of scholarly journals including the *Journal of Money, Credit, and Banking*, the *Journal of Banking & Finance,* and the *Journal of the Japanese and International Economies*.

Arito Ono is a professor at the Department of Banking and Corporate Finance, Faculty of Commerce, Chuo University, Japan. Prior to joining Chuo University in 2015, he was a senior economist at the Mizuho Research Institute, and a senior economist at the Institute for Monetary and Economic Studies, Bank of Japan (2009–2011). He has also been a member of several working groups at the Financial System Council, Financial Services Agency (2011–2015). His main fields of research are banking and corporate finance, and he has published papers in the *Journal of Money, Credit, and Banking*, the *Journal of Banking & Finance,* and the *Journal of Financial Stability*, among others. He received a B.A. in economics from the University of Tokyo and a Ph.D. in economics from Brown University.

Chapter 1
The Economics of Interfirm Networks: Main Issues

Tsutomu Watanabe, Iichiro Uesugi and Arito Ono

1.1 Introduction

Interfirm networks prevail in many facets of economic activity and are significantly influential across a range of economic phenomena from business cycles to knowledge spillovers. There is a growing concern in the policy arena relating to the vulnerability of such networks based on the casual observation that idiosyncratic shocks on firms can be amplified through interfirm connections and can lead to a systemic crisis. Typical examples are the manufacturing supply-chain networks in the automobile and electronics industries that propagated regionally concentrated shocks–such as the Great East Japan Earthquake and the floods in Thailand, both in 2011–into global ones. The recent global financial crisis has also shown that the failure of a large bank can have significant adverse effects on the economy as a whole via complex transaction networks.

There is growing interest among academic physicists and economists in network formation and functions. The standard economic model assumes that agents interact anonymously in centralized spot markets where transactions occur through independent decisions. However, many markets do not function as such, but rather involve many interactions through bilateral links. Hence, the notion of networks as a collection of nodes with links between them can be a useful tool for understanding a number of economic phenomena: there is an abundance of theoretical literature including Dutta and Jackson (2002), Jackson (2008), Goyal (2007), and Benhabib

T. Watanabe (✉)
Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
e-mail: watanabe@e.u-tokyo.ac.jp

I. Uesugi
Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, Japan
e-mail: iuesugi@ier.hit-u.ac.jp

A. Ono
Chuo University, 742-1 Higashinakano, Hachioji-shi, Tokyo, Japan
e-mail: a-ono@tamacc.chuo-u.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_1

et al. (2011) on this concept. However, there are a limited number of empirical studies on interfirm networks.

It is against this background that we instigated the research project titled "Designing industrial and financial networks to achieve sustainable economic growth" under Japan's Ministry of Education, Culture, Sports, Science and Technology program called "Promoting social science research aimed at solutions of near-future problems" in summer 2008. The main objective of the project is to examine the formation of interfirm networks and to investigate their impacts on a broad range of economic activities. The relevance of the research topic further increased in the face of subsequent massive shocks including the recent global financial crisis that erupted in late 2008 and the Great East Japan Earthquake that occurred in March 2011.

Having access to a comprehensive database on interfirm and firm-bank relationships was critical to the project success. For this purpose, we collaborated with the Teikoku Databank (TDB) Ltd.–the largest credit database company in Japan–and constructed a unique and massive transaction relationship database of approximately 400,000 Japanese firms. The TDB database provided details on firm attributes including their addresses (geographical location) and managerial performance and further providing information including their suppliers, customers, shareholders, and financial institutions. We augmented the database with variables from other sources including: information on firms' domestic and foreign investment activities extracted from government statistics, information on financial institutions' balance sheets obtained from other database companies, and information on the evolution of these networks based on our independent survey results. These database additions enabled us to examine a novel set of issues including the interaction between networks and economic agglomerations, and the impacts of interfirm transaction relationships and of firm–bank relationships on firms' investment behavior.

The pertinent research issues were split into three categories: (1) the structure and evolution of interfirm networks and their relationship with macroeconomic fluctuations, (2) the impact of interfirm networks on economic geography and on firm activities, and (3) the interactions between bank-firm relationships and firm behavior. The project led to a number of research articles on interfirm networks and relationships–including those compiled in this volume–that fall under one of the three identified category headings. Section 1.2 summarizes our research achievements by mainly focusing on the 11 papers collected in this volume and by selectively referring to some of our related research articles. Section 1.3 discusses several unresolved issues that are left for future research.

1.2 Summary of the Volume

1.2.1 Structure and Evolution of Interfirm Networks

The four chapters in Part I empirically examine the structure and evolution of interfirm networks in Japan through the lens of network analysis. In Chap. 2, Takayuki Mizuno, Wataru Souma, and Tsutomu Watanabe examine whether micro shocks to individual firms could diffuse widely to other firms through customer–supplier linkages and, ultimately, result in fluctuations in the economy as a whole. To this end, they start with a close look at the structure and evolution of customer–supplier networks in Japan from 2008 to 2012 using a unique dataset constructed from the TDB database. The authors present interesting empirical regularities in three aspects: (1) the cross-sectional attributes of the interfirm networks, (2) the evolution of interfirm networks, and (3) the interaction between the proximity of two firms in a network and the correlation of these firms' managerial performance. Regarding the cross-sectional characteristics of the interfirm networks, the authors find a non-uniform distribution of the customer and supplier links across firms, in that there are many small firms that have a few interfirm transaction relationships while there are a few "hub" firms that have many interfirm relationships. Moreover, the authors suggest that firms tend to have stronger incentives to acquire customers than they do to acquire suppliers. Finally, the authors find that the shortest path length for a given pair of firms is, on average, 4.3 links, suggesting that firms are closely interconnected presumably because of the existence of hub firms. Regarding the evolution of customer–supplier networks, the authors find that the switching of customers and suppliers is relatively rare and that there is a close interconnectedness among firms. This finding suggests that a shock to a single firm can be easily propagated to an entire network since there is little substitution among potential transaction counterparts. Regarding the association between network structure and the correlation of firm performance, the authors report that the correlation in annual sales growth between two firms is greater when the shortest link path between the two firms further shortens. This suggests that a non-negligible portion of firm sale fluctuations stems from the propagation of an idiosyncratic shock to other firms that are closely linked to it via customer–supplier chains.

In Chap. 3, Takashi Iino and Hiroshi Iyetomi analyze the community structure of interfirm networks in Japan. They visualize both an overall network structure and the network structure of eight specific industrial sectors and show that the structure of these networks is highly heterogeneous. For example, construction firm networks have a highly hierarchical structure (contractors, sub-contractors, sub-subcontractors, etc.), while manufacturing firm networks are tightly connected and form a dense cluster. These results contrast with the conventional wisdom that the manufacturing industry has highly hierarchical connections, and the findings merit further investigation.

The authors formally identify communities in Japanese interfirm networks using modularity maximization. A community is defined as a group of tightly connected nodes, with the links between them being sparse. The modularity measures the difference between the actual link density within a network (community) subset and its expected value: the optimized division of a network yields the highest modularity values. Using this methodology, Iino and Iyetomi extracted 118 communities from a large firm-level database. Focusing on the characteristics of the 10 largest communities in terms of the number of constituent firms, they find that each community is strongly associated with its constituent firms' characteristics with respect to regions and to industry sectors. The authors also detect sub-communities within the 10 largest communities, because modularity optimization often fails to identify small but important communities embedded within large communities. They show that the 10 largest sub-communities within a given community are also characterized by the constituent firms' areal and industrial attributes. For example, the largest sub-community of the largest community (machinery manufacturers and wholesalers) contains major electrical appliance manufacturers, while the second largest sub-community contains an automobile manufacturing cluster led by Toyota Motor Corporation and its group firms. Overall, modularity maximization can detect clusters of firms, including those of automobile and electronics firms that are frequently taken as typical examples of industry agglomerations. Finally, the authors investigate the structure of inter-community relationships by measuring their "distances" and the directional features of their transactions.

In Chap. 4, Iichiro Uesugi focuses on several regional industrial agglomerations and examines the structure and evolution of their interfirm networks from a variety of perspectives including: the nature of interfirm transaction relationships as well as the developments in such relationships over time; firm participation in network activities other than supplier–customer transactions; and interactions between interfirm transaction relationships and other relationship types. To investigate these issues, the author used TDB database information in conjunction with the results of a unique firm-level survey. A questionnaire was distributed to over 14,000 manufacturing firms located in Japan's three major industry agglomerations: the Keihin, Higashi-Osaka, and Hamamatsu areas. Regarding the development of interfirm transaction relationships over time, Uesugi finds that their numbers tended to decrease rather than increase in the three major industry agglomerations over the past 10 years, especially in the case of small firms and transaction relationships involving local customers and suppliers. This coincides with a substantial decline in the number of manufacturing firms located in these agglomerations and indicates that these industry agglomerations may have experienced negative feedback from the declining number of interfirm connections. Regarding the second issue, the author indicates that apart from transaction relationships, many firms had established other interfirm link types and participated in group activities including those in industry associations and local chambers of commerce. This suggests that regional proximity and shared industry interests are important determinants of firm links. Regarding the third issue, the author suggests that bank-lending attitudes appear to be affected by the links among local firms. Specifically, hub customer firms for local suppliers are more likely to have their loan applications approved than hub supplier firms for local customers. A possible interpretation of this result is that banks perceive the externalities associated with hub firms that purchase from local firms to be greater than those associated with hub firms that sell to local firms, and hence are more accommodative in their lending attitudes toward the former. This suggests that banks internalize the externality associated with hub customer firms when making a loan decision. This idea is formally examined in Ogura et al. (2014): an important contribution arising from our research project. The authors examine the lending decision by banks when such banks take the borrowing firm's interfirm transaction network structure into account.

1.2.2 Networks, Economic Geography, and Firm Activities

Part II examines how interfirm networks and relationships are associated with economic geography and firm activities. In Chap. 5, Kentaro Nakajima provides an overview of the relevant recent literature. The author summarizes the studies that use aggregated data and refers to their limitations, and then reiterates the importance of using disaggregated interfirm transaction data. The author reviews the literature that uses two data sources: the interfirm transaction relationship data from TDB and from Tokyo Shoko Research Ltd. (TSR); and data from the Commodity Flow Survey (CFS) undertaken jointly by the U.S. Bureau of Transportation Statistics and the U.S. Census Bureau. Nakajima then examines the impacts of interfirm transaction networks on the following three aspects: the geographical concentration of firms, the geographical propagation of shocks, and corporate decision making. One of the reviewed studies–Nakajima et al. (2012)–focuses on the relationship between interfirm transaction networks and the geographical agglomeration of firms and is a notable research contribution of our project. The authors propose a methodology to measure the extent of the localization of interfirm transaction relationships in the spirit of the pairwise distance approach of Duranton and Overman (2005). The authors thus detect the localization of transaction relationships at the firm level by setting up counterfactual transaction counterparts. This powerful tool provided by Nakajima et al. (2012) enables researchers to examine the relationship between interfirm transaction networks and industry agglomeration.

In Chap. 6, Gilles Duranton shows how the information on economic activity networks is used to define geographical agglomerations. The author focuses on the labor market and proposes a simple but robust methodology to define metropolitan areas by an iterative aggregation of spatial units using the information on commuting flows between them. Essentially, a spatial unit A is aggregated to another spatial unit B if the share of the workers who work in B among all those that reside in A is above a given threshold. A further spatial unit C may next be aggregated to the union of A and B if, similarly, it sends a fraction of its commuters greater than the same threshold to this newly formed unit even though it may not have been possible to aggregate C directly to either A or B. This process of aggregation repeats until no further unit can be aggregated. Aside from its simplicity, the methodology has two advantages: transparency that avoids possible political pressures on the definition of metropolitan areas, and robustness on the threshold level of the ratio of commuters from one municipality to the other. The author studies the case of Colombia to find several empirical regularities. Duranton acknowledges that the methodology could be used to define another set of metropolitan areas using information on interfirm transaction networks. The author did not undertake this because of data limitations in Colombia; however, other researchers may apply the proposed methodology to countries where complete interfirm transaction network data are available. They may also examine how economic agglomerations that use the interconnectedness of the labor markets differ from those that use the interconnectedness of the goods and services markets.

In Chap. 7, Hiroyuki Okamuro and Kenta Ikeuchi examine the formation of interfirm and firm–bank transaction networks by start-up firms. Most start-up firms are vulnerable to failure because of a lack of internal business resources. External resources such as business and financial networks are essential for the survival and growth of such firms. These networks both provide start-up firms with access to external business resources and signal their trustworthiness to third parties. Several studies have shown that entrepreneurial networks contribute to start-up performance; however, few have examined the formation of business and financial networks by start-up firms. Against this background, the authors empirically explore the determinants of interfirm network formation by start-up firms at the early stage of their life cycle, focusing on the founders' characteristics. Based on the TDB database, the empirical results show that lengthy industry experience of 10 years or more on the part of the founder has a significant positive impact on the size of both business and financial networks, while having a university education positively affects both the size and quality of business and financial networks. Surprisingly, no distinct differences are detected between the determinants of business and financial networks. Moreover, in another important project research contribution Okamuro also examines the impact of interfirm transaction relationships on corporate performance (Okamuro and Nishimura 2013).

In Chap. 8, Yukiko Saito details the geographical propagation of shocks through interfirm networks. The author focuses on a single massive natural disaster–the Great East Japan Earthquake that occurred in March 2011–and studies the extent of the impact among firms that were located outside the earthquake-hit areas. In addition to affecting the firms located in areas directly hit by the earthquake, there is abundant anecdotal evidence suggesting that the earthquake indirectly damaged firms located outside the hit areas through supply-chain disruptions; however, there is limited empirical evidence on this. There are two different, although not mutually exclusive, mechanisms through which shocks are propagated: (1) when shocks are transmitted among transaction partners located in the physical proximity of the affected firm and (2) when a limited number of hub firms propagate shocks to the entire economy through a large number of transaction relationships. Using interfirm transaction data from approximately 800,000 firms in Japan, the author examines the extent of the latter propagation mechanism by studying how firms in the unaffected areas are linked to those in the affected areas. The result indicates that most firms in the unaffected areas have indirect relations with firms in the affected areas. Overall, the findings in Chap. 8 highlight the importance of firms being closely linked to each other even if they are geographically distant and emphasize that regional hub firms play a key role in spreading localized shocks.

1.2.3 Bank-Firm Relationships and Firm Dynamics

Part III focuses on a specific bilateral network: bank-firm relationships.[1] In Chap. 9, Hans Degryse, Vasso Ioannidou, and Steven Ongena review the literature on bank-firm relationships, or relationship banking. Although financial contracts are intrinsically non-exclusive in the sense that borrowing firms cannot credibly commit to taking loans from just one bank and banks cannot prevent borrowers from taking credit from other lenders, they often engage in exclusive relationships. The chapter begins with a review of theoretical and empirical studies that provide rationales for exclusive bank-firm relationships. These studies argue that non-exclusive loan contracts would exacerbate the moral hazard incentives of borrowers. This in turn would worsen their access to credit and their credit terms from the initial banks that had the exclusive relationships. When non-exclusivity is pervasive, competing banks may fail to internalize the consequences of the future indebtedness of the borrowing firms through loans from other lenders, resulting in "overlending." Notably, however, engaging with multiple banks may also resolve exclusive relationship issues, and the costs and benefits of relationship banking to firms and banks are discussed. The benefits of relationship banking accruing to firms include, but are not limited to, increased credit availability, more flexibility in loan contract terms, and enhanced reputation. The costs of relationship banking for firms include higher lending rates associated with the "hold-up" problem that arises because of the informational advantage that the relationship bank holds over the competing banks. In times of crisis, relationship banks can smooth out loan contract terms when a firm is in temporary difficulty (idiosyncratic shock). Studies on the recent global financial crisis suggest that relationship banking may not follow its normal pattern when a shock is systemic and the banks themselves are damaged. The final section of Chap. 9 discusses how monetary and/or business conditions may affect the formation of bank-firm relationships.

In Chap. 10, Arito Ono, Hirofumi Uchida, Souichirou Kozuka, and Makoto Hazama provide a comprehensive overview of bank-firm relationships in Japan using a large dataset constructed in cooperation with TDB.[2] The chapter has two focuses: (1) it provides a detailed account of Japan's current "main bank" relationships and (2) it provides a unique and comprehensive description of the use of collateral in business financing in Japan. The chapter's authors find that the main bank relationships are stable: annually less than 1 % of firms switch their main banks. However, over 80 % of firms establish relationships with multiple banks, suggesting that there may not be an acute "hold-up" problem for most Japanese firms. The authors further find that main bank relationships are stronger in terms of deposit transactions than borrowing. This finding may imply that the traditional definition of Japanese main

[1] To date, the applications of network theory on finance mainly focus on issues related to the contagion among financial institutions, typically through interbank markets (Allen and Babus 2008). An exception to this is Ogura et al. (2014), which is a product of this research project.

[2] Using the same dataset, Uchida et al. (2015) provide a comprehensive overview of interfirm relationships in Japan, with emphasis on the use of trade credit.

banks used in many previous studies as those with the largest outstanding loan to firms is no longer valid. Real properties, the most typical form of collateral in Japan, are frequently used as collateral for multiple security interests, suggesting that junior liens are relatively common in Japan. The authors find that nearly 30 % of such properties are only secured by non-main banks: this casts doubt on the conventional view that the main banks possess senior security interests. Overall, the findings in Chap. 10 offer a unique representation of the current main bank relationships and provide a basis for further research on relationship banking in Japan, including that of Chap. 11.[3]

As the survey in Chap. 9 shows, relationship banking has benefits as well as costs. In Chap. 11, Kazuo Ogawa empirically examines this issue from the cash-holding behavior of Japanese firms based on a panel dataset in the 2000s. Specifically, the author investigates whether a stronger bank-firm relationship might economize firm cash holdings as these firms can rely more on liquidity provision by their main banks in difficult times (benefit of main bank relationships). An alternative hypothesis is that firms with stronger main bank relationships might be required to keep a larger amount of cash in the main banks' deposit accounts because of the "hold-up" problem (cost of main bank relationships). The author finds that firms that rely more on their main banks in terms of outstanding loans hold less cash. Additionally, the cash holdings of those firms with stronger main bank relationships are less sensitive to the changes in net working capital (transaction motive of cash holdings) and cash flow (precautionary motive of cash holdings). These findings support the hypothesis that stronger main bank relationships are beneficial to firms in economizing cash holdings. However, Ogawa also finds that firms pay a price for maintaining long-term stable relationships with the main bank in the form of higher effective borrowing rates that take into account the compensating balances to main banks. Whether the benefit of stronger relationships for firms in economizing cash holdings outweighs the associated costs in paying higher effective borrowing rates is an interesting question that is left for future research.

In Chap. 12, Kaoru Hosono and Daisuke Miyakawa review the extant empirical literature on how bank lending affects firm activities. As noted in Chap. 9, the formation, benefits, and costs of relationship banking may depend on the business cycle phase. Particularly, a financial crisis that inflicts severe damage to banks may propagate to the client firms that have close relationships with the damaged banks. There is much research on how shocks to banks affect their client firms' activities–such as survival/exit and investments–but many of these studies find it difficult to disentangle shocks to banks and shocks to firms. That is, although many studies found an empirical association between negative shocks to banks and a decline in their client firms' real activities, it is difficult for researchers to establish a causal link running from bank shocks to firm activities. The authors review the recent literature covering

[3] Employing the dataset of real properties collateral in Chap. 10, Ono et al. (2014) examine the evolution of loan-to-value (LTV) ratios during the bubble and post-bubble periods in Japan, and the ex-post performance of borrowers that obtained high LTV loans.

innovative identification strategies and classify such studies into three. First is the event-study approach that examines the stock market performance of firms that have relationships with failed banks. Second are studies that use geographical or economic borders: this is a promising line of research when shocks to banks and shocks to firms are isolated by such borders. The third classification is studies that use loan-level datasets that are also useful in identifying a supply or demand shock. For example, when a firm borrows from multiple banks, the relative differences in the loans from the different banks are likely to stem from a bank-specific loan supply shock because these loans originate from the same firm and thus share the same (firm-level) demand shock. The authors also review studies that use other identification strategies such as instrumental variables. Finally, Hosono and Miyakawa detail their personal studies that use a natural experiment approach–Hosono et al. (2012) and Miyakawa et al. (2014)–and that are products of the overall research project. Both studies examine firms that are not damaged by the natural disaster and consider that the same disaster that adversely affects their main banks is a purely exogenous loan supply shock. They find that damage to banks that were affected by Japan's Kobe Earthquake in January 1995 had significant negative impacts on their client firms' investments and exports that were located outside the earthquake-affected areas.

1.3 Issues for Future Research

The studies in this volume describe a number of important empirical results that contribute to the understanding of interfirm networks; however, some unresolved issues remain for future research to address.

The first issue relates to the quantitative impacts of the shocks propagated through networks. The analyses in Chaps. 2 and 8 show that the interfirm networks created through supplier–customer relationships in Japan exhibit a small-world structure in which a pair of firms are closely interconnected by a short path length. This suggests that a shock can propagate quickly into the entire economy. Analysis in Chap. 2 further shows that an interfirm network propagates small idiosyncratic shocks to aggregate shocks, suggesting the quantitative importance of the shock-amplifying mechanism through an interfirm network. In contrast, some studies provide evidence against the economic relevance of shock propagation through interfirm networks. For example, in a contribution to our research project, Hazama and Uesugi (2012) focus on the chain reaction of defaults in firm transaction networks in Japan and find that its extent is rather limited. Their empirical results suggest that banks may have acted as "deep pockets" and stopped the shock propagation; that is, banks extended credit to firms that had defaulting customers and thus helped the firms act against their own possible default.

In addition to banks, there are a number of other factors that potentially affect the extent that idiosyncratic shocks propagate into massive macro shocks including: the substitutability among trading counterparts, the direction of the flow of goods and services between firms, and the commitment by the government to help financially distressed firms in interfirm networks. To definitely conclude whether the propagation of shocks through interfirm networks impacts on the entire economy, such factors should be considered in conjunction with the results on the quantitative impacts of shocks through the interfirm transaction networks.

The second issue needing further research relates to the persistence of shocks propagated through interfirm networks. On one hand, there is evidence–including Rogoff and Reinhart (2009)–that a financial crisis has a long-lasting impact on the real economic activities, suggesting that an idiosyncratic shock can persistently affect an entire economy. On the other hand, anecdotal evidence on the aftermath of supply-chain disruptions after the Great East Japan Earthquake indicates that the effect was relatively short-lived.[4] Additionally, the empirical studies by Hosono et al. (2012) and Miyakawa et al. (2014) introduced in Chap. 12 show that the negative impact of the Kobe Earthquake on firm investments and exports disappeared after approximately 2 years. However, there is insufficient empirical evidence on this topic to make a solid inference and further studies are needed.

Third, the chapters in this volume provide abundant evidence for the existence of either positive (Chaps. 2 and 7) or negative (Chaps 2, 8, and 12) network effects. However, the extent to which firms formulate interfirm networks taking these pros and cons into account is less understood. In this regard, the literature on relationship banking (details on some studies are included in Part III) provides a useful starting point for future research. Although the literature focuses on a specific bilateral network, there are many studies that examine: the intrinsic features of firms and banks whose relationship banking benefits outweigh the costs; how firms and banks exit from existing relationships when their *net* benefits turn negative; and with whom firms and banks formulate new relationships. A synthetic analysis of the benefits and costs of interfirm networks is a promising line of research for future studies.

References

Allen, F., & Babus, A. (2008). Networks in finance. Wharton Financial Institutions Center Working Paper, No. 08-07.

Benhabib, J., Bisin, A., & Jackson, M. O. (2011). *Handbook of social economics Vol. 1A*. Amsterdam: Elsevier (North Holland Publishing).

Duranton, G., & Overman, H. (2005). Testing for localization using micro-geographic data. *Review of Economic Studies, 72*(4), 1077–1106.

Dutta, D., & Jackson, M. O. (2002). *Networks and groups*. Berlin: Springer.

Goyal, S. (2007). *Connections: An introduction to the economics of networks*. Princeton: Princeton University Press.

Hazama, M., & Uesugi, I. (2012). Measuring the systemic risk in interfirm transaction networks. HIT-TDB Network Center Working Paper, No. 20.

[4] Todo, Nakajima, and Matous (2015) evidence that the firms that have trading partners located outside the earthquake-affected area resumed their operations earlier.

Hosono, K., Miyakawa, D., Uchino, T., Hazama, M., Ono, A., Uchida, H., & Uesugi, I. (2012). Natural disasters, damage to banks, and firm investment. HIT-TDB Network Center Working Paper, No. 18.

Jackson, M. O. (2008). *Social and economic networks*. Princeton: Princeton University Press.

Miyakawa, D., Hosono, K., Uchino, T., Ono, A., Uchida, H., & Uesugi, I. (2014). Financial shocks and firm exports: A natural experiment approach with a massive earthquake. RIETI Discussion Paper Series, 14-E-10.

Nakajima, K., Saito, Y., & Uesugi, I. (2012). Localization of interfirm transaction relationships and industry agglomeration. RIETI Discussion Paper Series, 12-E-23.

Ogura, Y., Okui, R., & Saito, Y. (2014). Network-motivated lending decision. Mimeo.

Okamuro, H., & Nishimura, J. (2013). Knowledge and rent spillovers through Government-Sponsored R & D Consortia. *HIT-TDB Network Center Working Paper*, No. 24.

Ono, A., Uchida, H., Udell, G. F., & Uesugi, I. (2014). Lending pro-cyclicality and macro-prudential policy: Evidence from Japanese LTV ratios. TCER Working Paper Series, E-70.

Rogoff, K. S., & Reinhart, C. M. (2009). *This time is different: Eight centuries of financial folly*. Princeton: Princeton University Press.

Todo, Y., Nakajima, K., & Matous, P. (2015). How do supply chain networks affect the resilience of firms to natural disasters? Evidence from the Great East Japan earthquake. *Journal of Regional Science, 55*(2), 209–229.

Uchida, H., Ono, A., Kozuka, S., Hazama, M., & Uesugi, I. (2015). *Interfirm Relationships and Trade Credit in Japan: Evidence from Micro-Data*. Springer Briefs in Economics, Tokyo: Springer.

Tsutomu Watanabe has been a professor of economics at the Graduate School of Economics, The University of Tokyo, since October 2011. He was a professor at Hitotsubashi University (1999–2011) and a senior economist at the Bank of Japan (1982–1999). He has held visiting positions at various universities, including Kyoto University, Bocconi University, and Columbia University. He received his Ph.D. in economics from Harvard University in 1992 and did his undergraduate work at The University of Tokyo. Watanabe's main research area is monetary policy and inflation dynamics. He is known for his series of papers on monetary policy when nominal interest rates are bounded at zero; in particular, his paper on the optimal monetary policy at the zero lower bound has been widely recognized as the first paper to provide a simple description of the liquidity trap and characterize the optimal policy response to it in a setting of dynamic stochastic general equilibrium. He is an author of many books and more than 40 academic journal articles on monetary policy, inflation dynamics, and international finance.

Iichiro Uesugi is a professor at the Institute of Economic Research of Hitotsubashi University, Japan. He received his M.A. and Ph.D. in economics in 2000 from the University of California, San Diego. Prior to becoming a member of the faculty of Hitotsubashi University in 2011, he joined the Ministry of Economy, Trade and Industry (METI) in 1993 and worked in the Economic Policy Unit of the Minister's Secretariat, Macroeconomic Affairs Division; the Planning Division of the Coal Department in the Agency of Natural Resources and Energy; the Research Division of the Small and Medium Enterprises Agency; and the Research Institute of Economy, Trade and Industry. His research interests focus on firm dynamics, financial intermediation, and the Japanese economy. His research has been published in a number of scholarly journals including the *Journal of Money, Credit, and Banking*, the *Journal of Banking & Finance,* and the *Journal of the Japanese and International Economies*.

Arito Ono is a professor at the Department of Banking and Corporate Finance, Faculty of Commerce, Chuo University, Japan. Prior to joining Chuo University in 2015, he was a senior economist at the Mizuho Research Institute, and a senior economist at the Institute for Monetary and Economic Studies, Bank of Japan (2009–2011). He has also been a member of several working groups at the

Financial System Council, Financial Services Agency (2011–2015). His main fields of research are banking and corporate finance, and he has published papers in the *Journal of Money, Credit, and Banking, the Journal of Banking & Finance*, and the *Journal of Financial Stability*, among others. He received a B.A. in economics from the University of Tokyo and a Ph.D. in economics from Brown University.

Part I
Structure and Evolution of Interfirm Networks

Chapter 2
Buyer-Supplier Networks and Aggregate Volatility

Takayuki Mizuno, Wataru Souma and Tsutomu Watanabe

Abstract This chapter investigates the structure and evolution of customer–supplier networks in Japan using a unique dataset that contains information on customer and supplier linkages for over 500,000 incorporated non-financial firms for the 5 years from 2008 to 2012. We find, first, that the number of customer links is unequal across firms: the customer link distribution has a power-law tail with an exponent of unity (i.e., it follows Zipf's law). We interpret this as implying that competition among firms to acquire new customers yields winners that attract a large number of customers, as well as losers that end up with fewer customers. We also show that the shortest path length for any pair of firms is, on average, 4.3 links. Second, we find that link switching is relatively rare. Our estimates indicate that 92 % of customer links and 93 % of supplier links survive each year. Third and finally, we find that firm growth rates tend to be more highly correlated the closer two firms are to each other in a customer–supplier network (i.e., the smaller is the shortest path length for the

This paper is an extended version of Mizuno et al. (2014).

T. Mizuno (✉)
National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan
e-mail: mizuno@nii.ac.jp

Department of Informatics, The Graduate University for Advanced Studies, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

PRESTO, Japan Science and Technology Agency, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

The Canon Institute for Global Studies, 5-1 Marunouchi 1-chome, Chiyoda-ku, Tokyo, Japan

W. Souma
College of Science and Technology, Nihon University, 7-24-1 Narashinodai, Funabashi, Chiba, Japan
e-mail: soumawataru@nihon-u.ac.jp

T. Watanabe
Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
e-mail: watanabe@e.u-tokyo.ac.jp

The Canon Institute for Global Studies, 5-1 Marunouchi 1-chome, Chiyoda-ku, Tokyo, Japan

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_2

two firms). This suggests that a non-negligible portion of firm growth fluctuations stem from the propagation of microeconomic shocks—shocks that affect a specific firm—through the customer–supplier chains.

Keywords Buyer-supplier network · Aggregate volatility · Input–output analysis · Power-law distribution · PageRank

2.1 Introduction

Firms in a modern economy tend to be closely interconnected, particularly those in the manufacturing sector. Firms typically rely on the delivery of materials or intermediate products from their suppliers to produce their own products that in turn are delivered to downstream firms. Two recent episodes vividly illustrate how closely firms are interconnected. The first is the earthquake and tsunami that hit the Tohoku region in the northeastern part of Japan on March 11, 2011, resulting in significant human and physical damage to that region. However, the economic damage was not restricted to the geographical region and spread in an unanticipated manner to other parts of Japan through the disruption of supply chains. For example, vehicle production by Japanese automakers that were physically removed from the affected areas was either halted or reduced because of a shortage of auto part supplies from firms located in the affected areas. This shock even spread across borders and led to a substantial decline in North American vehicle production.[1] The second episode is the recent financial turmoil triggered by the subprime mortgage crisis in the United States. This adverse shock originally stemmed from the so-called toxic assets on the balance sheets of U.S. financial institutions and led to the failure of these institutions. The shock was transmitted beyond entities that had direct business dealings with the collapsed financial institutions to those that seemed to have no relationship with them, resulting in an upheaval that affected financial institutions around the world.

These two episodes show that both national economies and the global economy are subject to the risk of a chain reaction in product disruptions through customer–supplier linkages. Such risk is particularly high when the linkage structure in the economy is dominated by a few hub firms whose products are supplied to many other firms as input. Importantly, supply-chain disruptions are more serious when there are no close substitutes to the hub firms, at least in the short run. Motivated at least partly by these episodes, some recent studies in macroeconomics have sought to develop theoretical production chain models that extend input-output analysis—which dates back to the seminal work by Wassily Leontief published in the 1930s

[1] For example, the U.S. Federal Reserve chairman Ben Bernanke stated in the aftermath of the earthquake: "U.S. economic growth so far this year looks to have been somewhat slower than expected. Aggregate output increased at only 1.8 % at an annual rate in the first quarter, and supply-chain disruptions associated with the earthquake and tsunami in Japan are hampering economic activity this quarter." (Speech at the International Monetary Conference, Atlanta, Georgia, U.S. on June 7, 2011).

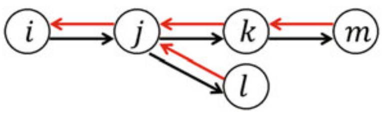


Fig. 2.1 An illustration of customer–supplier network consisting of five firms. The *red arrows* in the figure indicate the flow of money, while the *black arrows* indicate the flow of products each firm produces. For example, firm j purchased something from firm i and sells something to firms k and l

(Leontief 1936)—to identify conditions under which microeconomic shocks (idiosyncratic shocks to individual firms) can propagate to the rest of the economy through production chains, leading to fluctuations in production at the macro level (Acemoglu et al. 2012, 2013b; Carvalho 2010, 2012; Stella 2013; Kelly et al. 2013). Policymakers are also aware of the need to prepare for the propagation of adverse shocks through production chains.[2]

The present study seeks to provide empirical evidence on the structure and evolution of customer-supplier networks in Japan using a unique dataset that contains information on customer and supplier linkages for over 500,000 incorporated non-financial firms for the 5-year period from 2008–2012. This dataset provides the customer and supplier lists for each firm. We use these lists to produce a customer-supplier network. To illustrate this, Fig. 2.1 shows a simple example consisting of five firms. The red arrows in the figure indicate the flow of money, while the black arrows indicate the flow of products that each firm produces. Firm j purchases something from firm i (its supplier) and sells something to firms k and l (its customers). Moreover, firm k, which purchases something from firm j, sells to firm m. Note that firm k is both a customer (it buys from firm j) and a supplier (it sells to firm m).

The rest of the paper is organized as follows. Section 2.2 compares two important customer–supplier network models—Leontief and PageRank—showing that both models are equivalent under some assumptions. Section 2.3 provides a detailed description of the dataset used, while Section 2.4 examines the basic structure of customer–supplier networks, focusing particularly on how closely firms are interconnected. Section 2.5 investigates how customer–supplier networks evolve over time. Section 2.6 empirically evaluates the extent that firm sales and growth are affected by the propagation of idiosyncratic shocks through production chains. Finally, Section 2.7 concludes the paper.

[2] The study of networks as phenomena that deserve analysis goes back to the small-world network model by Watts (Watts and Strogatz 1998) and has gained popularity in a variety of scientific disciplines including statistical physics, computer science, biology, and sociology. The methodology developed in those disciplines has been introduced into economics only relatively recently (Jackson 2010; Goyal 2012). However, it has produced important contributions on bank-firm relationships (Souma 2003), cross shareholdings (Garlaschelli et al. 2005), supply chains (Atalay et al. 2011; Saito et al. 2007; Ohnishi et al. 2010; Takayasu et al. 2008; Fujiwara and Aoyama 2010; Watanabe et al. 2012), systemic risks in financial markets (Battiston et al. 2007; Acemoglu et al. 2013a), and international trade (Garlaschelli and Loffredo 2005; Garlaschelli and Loffredo 2004; Di Giovanni and Levchenko 2010).

2.2 Equivalence of Leontief and PageRank Models

The initial concept of interfirm networks originated in the 1930s (Leontief 1936), although the nodes used in the Leontief analysis at that time are economic sectors (i.e., industries) rather than individual firms. Let us denote the number of firms in an economy by N and the sales vector associated with those firms by x that is a column vector with the sales of firm i on the ith row. We normalize x so that $x'\mathbf{1} = 1$ holds. The input-output structure of the economy is represented by an $N \times N$ matrix A with a_{ij} as an (i, j) element. The element a_{ij} denotes the share of input j (i.e., the commodity produced by firm j) in the total intermediate input use of firm i. Market clearing conditions are given by

$$x = (1 - \alpha)A'x + f \tag{2.1}$$

where f is a column vector representing the final demand to individual firms 'and α is the share of value added to gross sales. The first and second terms on the right-hand side represent the intermediate and final demands. Two new assumptions are now introduced. The first assumption relates to the final demand vector that is given by

$$f = \frac{\alpha}{N}\mathbf{1} \tag{2.2}$$

That is, the final demand is equal across the firms. Second, we assume that the a_{ij} is equal to the reciprocal of the total number of suppliers to firm i if $a_{ij} > 0$ and zero otherwise. This means that the supply links to firm i are of the same thickness. A new input-output matrix defined in this way is denoted by $\tilde{A}$. Given these assumptions, Eq. (2.1) changes to

$$\tilde{x} = (1 - \alpha)\tilde{A}'\tilde{x} + \frac{\alpha}{N}\mathbf{1} \tag{2.3}$$

Notably, the column vector $\tilde{x}$ in Eq. (2.3) is a PageRank vector (Brin and Page 1998). PageRank is an algorithm used by Google Search to rank websites in their search engine results. Equation (2.3) shows that the input-output model invented by Leontief in the 1930s is closely connected to the basic idea of PageRank.

Based on Eq. (2.3), Acemoglu et al. (2012) investigates how an economy's value added, the log of which is denoted by y, is affected by idiosyncratic shocks to individual firms in the economy. Denoting an idiosyncratic shock to firm i by ν_i (which is assumed to be i.i.d. with mean zero and variance σ^2) and the corresponding column vector by ν, we have $y = \tilde{x}'\nu$. If the distribution of PageRank across firms (i.e., $\tilde{x}_i$ for $i = 1, \ldots, N$) follows a uniform distribution (i.e., $\tilde{x}_i = 1/N$), then we have $\sqrt{\mathrm{var}(y)} = \sigma/\sqrt{N}$, implying that the standard deviation of y converges to zero as $N \to \infty$ at the rate $\sqrt{N}$. Generally, the central limit theorem guarantees that the standard deviation of y decays at the rate $\sqrt{N}$ if the PageRank distribution is sufficiently close to a uniform distribution. This implies that idiosyncratic shocks to individual firms would not translate into aggregate shocks because idiosyncratic shocks quickly cancel each other out as the number of firms increases (Dupor 1999).

However, this does not hold if PageRank is substantially unequal across firms (Acemoglu et al. 2012). Specifically, if the PageRank distribution has a power-law tail (i.e., $\Pr(x_i > x) \propto x^{-\zeta}$, where ζ is a power-law exponent with ζ between 1 and 2), then we have $\sqrt{\text{var}(y)} = \sigma/N^{1-1/\zeta}$. This means that the standard deviation of y decays at a rate slower than $\sqrt{N}$, implying that idiosyncratic shocks do not cancel each other out as quickly as implied by the central limit theorem. Hence idiosyncratic shocks to firms with very large PageRank may have a substantial impact on y. Typically, firms with large PageRank are hub firms that have a large number of trade partners. Idiosyncratic shocks to those hub firms spread to other firms through customer–supplier linkages, leading to a cascade phenomenon.

Notably, this argument is based on the two assumptions regarding final demand and the input-output matrix that may not actually hold in the data. However, one can make a similar argument by replacing $y = \tilde{x}'\nu$ with $y = x'\nu$. Gabaix (2011) shows that var(y) converges to zero as $N \to \infty$ at the rate $\sqrt{N}$ if the x_i's are uniformly distributed (or close to it). Conversely, var(y) decays at a slower rate if the distribution of x_i is heavy tailed, implying again that idiosyncratic shocks to individual firms would translate into macro shocks. This is referred to by Gabaix (2011) as granular hypothesis. It differs importantly from the cascade hypothesis proposed by Acemoglu et al. (2012) in that firms with large x_i may not necessarily be highly connected. For example, the large x_i of those firms may come from final demand rather than intermediate demand.

These two hypotheses have different implications on how policy makers should act to mitigate fluctuations in y. The granular hypothesis considers that fluctuations in y come from firms with large sales, so that it is important to mitigate idiosyncratic shocks to those firms. The "too big to fail" principle, often discussed in the context of preventing the failures of large financial institutions, is an example of such an action. The cascade hypothesis, however, implies that how closely a firm is connected to other firms through its customer–supplier linkages is vital rather than the firm size. This corresponds to the idea of "too interconnected to fail" that has been discussed by Markose et al. (2012) among others in the context of the recent financial crisis.

The cascade hypothesis has two testable implications. The first implication is that the number of trade links is highly unequal across firms. Particularly, the number of customer links for a firm, which is closely related to its PageRank, must be highly unequal and its distribution must have a heavy upper tail. Second, the cascade hypothesis implies that firm growth rates should be more highly correlated the closer two firms are to each other in a customer–supplier network. We test both implications using a dataset that contains information on customer and supplier linkages for over 500,000 incorporated non-financial firms.

2.3 Data

The dataset we use is jointly compiled by Teikoku Databank, Ltd. (TDB), one of the largest business database companies in Japan, and the HIT-TDB project of Hitotsubashi University. The dataset mainly provides information related to corporate bankruptcies and credit ratings and covers about 1.3 million incorporated non-financial firms. Since the number of corporations in Japan in 2006 (as reported in the 2006 Establishment and Enterprise Census) was 1493 million, our dataset covers about 90 % of all incorporated firms in Japan. TDB collects various kinds of information from these firms, including annual or more frequent financial statement data.

Two types of information on customer–supplier relationships are recorded in this dataset. First, the dataset contains information on the number of three types of relationships a firm has with other firms, namely relationships with customers (i.e., firms to which a firm sells its products), suppliers (i.e., firms from which a firm purchases raw materials and intermediate products), and owners (i.e., firms by which a firm is owned). Since in this paper we focus on customer–supplier relationships, we mainly use information on customer and supplier linkages. We denote the total number of firm i's customer links by N_i^C and the total number of supplier links by N_i^S. Second, the dataset lists the firms with which a firm has links (i.e., customers or suppliers to the firm) with their identification codes. However, the list is not exhaustive and its length cannot exceed 60 firms. This means that for smaller firms with fewer than 60 partners all of their partners are listed, but for large firms with more partners only the 60 most important ones are listed. In all cases, transaction partners are listed in descending order of importance based on the transaction volume.

Table 2.1 presents descriptive statistics on customer and supplier linkages. All statistics in the table are calculated using the total number of linkages, that is, N_i^C and N_i^S. Note that the table provides linkage information for five different years (i.e., 2008, 2009, 2010, 2011, and 2012), allowing us to investigate not only the structure of customer–supplier networks at a particular point in time but also their evolution. The sample mean for the number of customer links per firm is about 340 each year, and the median for the number of customer links per firm is 50, which is about one seventh of the mean, implying that the customer link distribution is not symmetric, but is substantially skewed to the right. In fact, the maximum number of customer links in 2012 was 95,512, which is far greater than the mean or the median, given that the standard deviation is only 2053. Turning to the number of supplier links, the sample mean is about 60 each year, which is much smaller than the number of customer links. A typical firm has six times as many customer links as supplier links. The median number of supplier links per firm is 20, implying again that the distribution for the number of supplier links is not symmetric but is skewed to the right. The maximum number of supplier links per firm is also much greater than the mean or the median.

To investigate the structure of customer–supplier networks and their evolution over time, we use the list of firms linked to a firm with their identification codes. As mentioned, the list is not exhaustive, so that, as far as large firms are concerned, links with less important partners are not recorded. The number of customers and suppliers in the list is 6.7 and 6.4 for a typical firm, which is much smaller than the means of the *total* number of customer and supplier links presented in Table 2.1. We augment the customer/supplier lists as follows. We first identify firm A as a supplier of firm B using the *customer* list of firm A, thereby producing an augmented supplier list of firm B. We add up the number of customer links originally shown in the customer list of a

Table 2.1 Number of customer and supplier links per firm

Customer links	2008	2009	2010	2011	2012
Number of firms	160,508	155,806	144,006	142,931	145,317
Number of links per firm					
Mean	339	343	341	340	339
Median	50	50	50	50	50
Std. Dev.	2107	2090	2015	2022	2053
Max.	90,200	90,504	90,000	90,000	95,512
Min.	0	0	0	0	0
Supplier links	2008	2009	2010	2011	2012
Number of firms	215,562	208,459	192,111	189,493	193,045
Number of links per firm					
Mean	56	58	61	62	61
Median	20	20	20	20	20
Std. Dev.	281	314	368	332	351
Max.	52,100	55,100	70,000	70,000	70,000
Min.	0	0	0	0	0

firm and the number of customer links identified in this way, and denote the sum by $\tilde{N}_i^C$. Similarly, we use the supplier lists of firms to produce augmented customer lists and define $\tilde{N}_i^S$. This kind of "reverse lookup" method has been applied to different datasets in previous studies on interfirm relationships, including Saito et al. (2007), Fujiwara and Aoyama (2010), Takayasu et al. (2008). Comparing N_i^C and $\tilde{N}_i^C$, we observe a relationship of the following form:

$$\left\langle \tilde{N}^C \mid N^C = n \right\rangle \propto n^{0.83} \qquad \text{for} \qquad 20 \le n \le 10000. \tag{2.4}$$

where $\langle \tilde{N}^C \mid N^C = n \rangle$ represents the mean of $\tilde{N}_i^C$ across i given that the total (true) number of customer links, N^C, for those firms is equal to n. Interestingly, the power exponent of n is smaller than unity, implying that for firms with a large number of customers the augmented list still does not capture the true number of customers. The example of a firm leasing vending machines to other firms explains why. This firm has a very large number of customer firms, but because vending machines are not regarded as a key input to production by most customer firms, they do not include the leasing firm in their list of suppliers. In this case, $\tilde{N}^C$ for the leasing firm is much smaller than N^C.

Turning to supplier lists, we have

$$\left\langle \tilde{N}^S \mid N^S = n \right\rangle \propto n^{1.19} \qquad \text{for} \qquad 10 \le n \le 1000. \tag{2.5}$$

indicating that the exponent of n is now greater than unity, which means that $\tilde{N}^S$ more than doubles when N^S doubles, and in this sense $\tilde{N}^S$ overestimates N^S. A likely

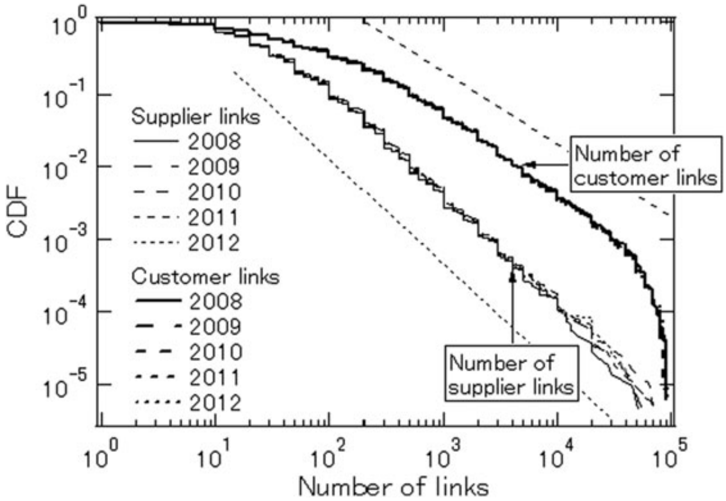


Fig. 2.2 Cumulative distributions of customer and supplier links in 2008–2012. The horizontal axis represents the *total* number of links, i.e., N^C and N^S, while the vertical axis represents the corresponding cumulative densities. The dotted straight lines are reference lines with a slope of -1 and -1.5 respectively. The number of firms used in this calculation is shown in Table 2.1

reason is that small suppliers to a prestigious firm with a large number of suppliers will include the prestigious firm in their customer list reported to TDB, since the prestigious firm is regarded as a key constituent of their customer base. However, this effect will be weak or absent if a customer firm is not that prestigious, which makes the exponent of n in Eq. (2.5) greater than unity.

2.4 The Structure of Customer–Supplier Networks

2.4.1 Unequal Links Across Firms

The number of links is unequal across firms with regard to both customer and supplier linkages, as we saw in Table 2.1. One may wonder how unequal it is across firms and whether the degree of inequality differs between customer and supplier linkages. To address these questions, we show in Fig. 2.2 the cumulative distribution functions (CDFs) of links across firms. The horizontal axis represents the number of links, while the vertical axis shows the corresponding cumulative densities. The horizontal and vertical axes are both in logarithm. For example, the number on the vertical axis corresponding to 10^2 on the horizontal axis is about 10^{-1} for supplier linkages, indicating that firms with more than 10^2 supplier links account for one tenth of all firms. The figure shows the CDFs for the customer and supplier linkages for each of our five observation years (2008, 2009, 2010, 2011, and 2012).

Given that the mean for the logarithm of the number of customer links is 1.72 and the corresponding standard deviation is 0.783, a number like 5000 links deviates from the mean by more than 2.52σ, and a number like 50,000 links deviates by more than 3.80σ. If the number of customer links is lognormally distributed, the cumulative probabilities corresponding to 5000 and 50,000 links are 0.0058 and 0.000072, which is much lower than the probabilities that we actually observe, indicating that the number of customer links has a heavier upper tail than a lognormal distribution.

The CDFs of customer links show a linear relationship between the log of the number of links and the log of the corresponding cumulative probability for the number of links within the range of 80–50,000. The slope is around -1 and is not significantly different from this value in each of the 5 years, that is,

$$P_>(N^C) \propto \frac{1}{N^C} \qquad \text{for} \qquad 80 \le N^C \le 50000 \tag{2.6}$$

where $P_>(N^C)$ represents the probability that the number of customer links exceeds a certain value. Equation (2.6) shows that N^C follows a power-law distribution and, more importantly, that its exponent is very close to unity. Power-law distributions with exponent 1 are found in various economic phenomena, including the distribution of city sizes, asset price changes, and firm sizes, a phenomenon referred to as Zipf's law. Most importantly, as shown by previous studies (Axtell 2001), firm sales follows Zipf's law, suggesting that the sales of a firm are related to the number of customers the firm has. We will come back to this issue in Section 2.5.

Turning to the number of supplier links, we again find a linear relationship between the log of the number of supplier links and the log of the corresponding cumulative density, indicating that the number of supplier links also follows a power-law distribution. However, the slope of the linear relationship is much larger than that in the case of customer links, implying that the tail part of the supply link distribution is less fat than that of the customer link distribution. The slope associated with supplier linkages is about -1.5, so that the CDFs for the number of supplier links can be characterized by

$$P_>(N^S) \propto \left(\frac{1}{N^S}\right)^{1.5} \qquad \text{for} \qquad N^S \ge 30. \tag{2.7}$$

Since the power-law exponent in this case exceeds unity, Zipf's law does not hold. Note that the power-law exponent ζ is related to the Gini coefficient, G, in the form $G = 1/(2\zeta - 1)$. Therefore, the fact that the power-law exponent is larger for supplier linkages than for customer linkages implies that the Gini coefficient is smaller for supplier linkages and that, therefore, the number of supplier links across firms is less unequal than the number of customer links.

What explains this result? As emphasized in the recent literature on customer search models (Luttmer 2006; Gourio and Rudanko 2011), firms spend substantial resources on marketing to acquire as many customers as possible in order to increase their sales and profits. Such competition among firms produces winners with a large number of customers as well as losers with a small number of customers, resulting

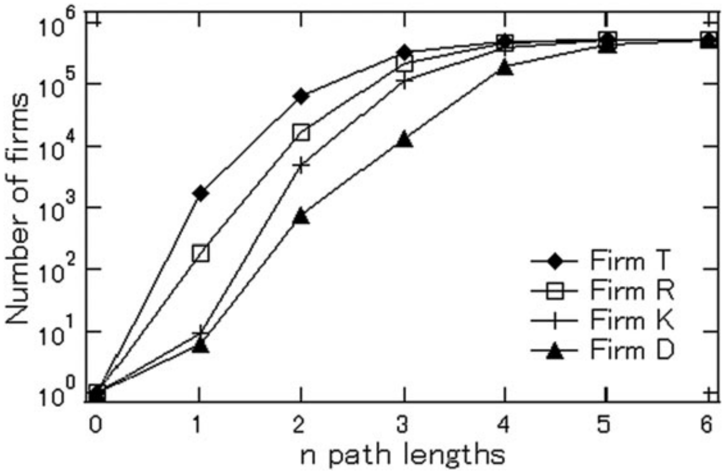


Fig. 2.3 Number of firms connected to a particular firm by *n* path lengths. Firms *T*, *R*, *K*, and *D* are randomly picked from the sample, which consists of all the firms on the augmented customer/supplier lists

in huge inequality in the number of customers. In contrast, with regard to supplier linkages, firms have little incentive to increase their number of suppliers because it is not necessarily profitable to buy materials and intermediate products from more suppliers. It may even be the case that purchases from more suppliers increase the associated costs (e.g., shipping costs) and therefore reduce profits. Therefore, because firms do not compete to have as many suppliers as possible, the extent of inequality is not as high as that with regard to the number of customers.

2.4.2 How Closely are Firms Interconnected?

To investigate how closely firms are interconnected, we use the augmented customer/supplier lists of partners mentioned in Section 2.2 for the set of firms whose identification codes are listed in the customer and/or supplier lists of the other firms. The number of firms that appear in the augmented lists is about 500,000.[3] Specifically, we randomly pick four firms (Firms T, R, K, and D) to examine the number of firms connected to a particular firm by one, two, three, or more path lengths. The result is shown in Fig. 2.3. Firm T is connected to about 1700 firms by one path length, but it is connected to more than 60,000 firms by two path lengths. The corresponding number for four path lengths increases to 503,796, which is only slightly

[3] The number of firms in the augmented lists is 552,145 for 2008, 541,816 for 2009, 518,565 for 2010, 520,087 for 2011, and 525,836 for 2012.

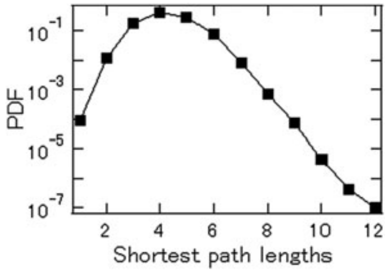


Fig. 2.4 Distribution of the shortest path lengths for all pairs of firms. We pick firms that are on the augmented customer/supplier lists for each year in 2008–2011 *and* whose sales data are available for 1980–2009, the number of which is 134,067. We calculate the shortest path length for every pair of firms. There are 17.9 billion pairs

less than the total number of listed firms. Thus, firm T is connected to almost all the firms by four path lengths or less.

The fact that firm T is connected to about 1700 firms by one path length, which is much larger than the sample average presented in Table 2.1, suggests that it is extremely large. Given firm T's size and the fact that it is connected to about 1700 firms by one path length, it may not be very surprising to find that it is connected to almost all the other firms by less than four path lengths. However, a more surprising case is Firm D, which is connected to only ten firms by one path length and, in fact, is very small with fewer than ten workers. Nevertheless, the number of firms to which Firm D is connected is 746 for two path lengths, 13,519 for three path lengths, 196,799 for four path lengths, and 446,019 for five path lengths. Surprisingly, even a small firm like Firm D is connected to almost all the listed firms by five path lengths or less.

We pick 130,000 firms that are on the customer/supplier lists for every year in 2008–2012,[4] and then calculate the shortest path lengths for every pair of firms. There are about 17.9 billion pairs and we find that 99.6 % of all pairs are connected, but 0.4 % cannot be connected regardless how long the path lengths are. Figure 2.4 shows the distribution of the shortest path lengths for those connected pairs. The mode of the distribution is four path lengths, and about 61.7 % of the pairs are connected by four path lengths or less. Note that a similar feature a customer–supplier network is reported by Ohnishi et al. (2010), Takayasu et al. (2008) using a different dataset. Also note that previous studies including (Fujiwara and Aoyama 2010) apply the technique of community analysis to customer–supplier networks to find that path length tends to be shorter between firms belonging to the same industry or located in the same region.

[4] More specifically, we pick 134,067 firms that are on the augmented customer/supplier lists for each year in 2008–2011 *and* whose sales data are available for 1980–2009. We will focus on the same set of firms in the analysis in Section 2.5.

Why are firms so closely interconnected? It is important to recall that the number of links, in the case of both customer and supplier linkages, follows a fat-tailed distribution. This indicates that there are some (although not many) firms with an extremely large number of links. The presence of such "hub" firms implies that even a small firm, like Firm D in Fig. 2.3, is able to be connected to a large number of firms through these hub firms; that is, once a small firm finds a path reaching one of the hub firms (probably via several steps), it is then connected to the large number of firms to which the hub firm is linked. This kind of small-world phenomenon can be found for various economic and social networks (Jackson 2010; Goyal 2012).

2.5 The Evolution of Customer–Supplier Networks

A distinctive feature of our dataset is that it records information on linkages for five different years, allowing us to investigate not only the structure of customer–supplier networks at a particular point in time, but also their evolution over time. Some firms continue to buy from the same suppliers and sell to the same customers for a long period. However, other firms change their partners quite often. The duration of customer–supplier relationships influences how shocks are transmitted through the network. Suppose that a firm is hit by an adverse shock, and the firm reduces its production. If the relationships are all fixed and the network therefore is highly stable, then the shock to that firm spreads to downstream firms, which are also forced to reduce production. However, if relationships are flexible in the sense that firms can change their customers/suppliers easily (i.e., without incurring any large costs), downstream firms can easily establish new supplier links and thereby keep the shock from spreading.

To see how quickly customer and supplier networks evolve over time, we present in Table 2.2 some statistics related to the turnover of customers and supplier links. Specifically, we identify customer links that appear on the augmented customer list of a firm in 2008 but not in 2009 and count them as link exits. Similarly, we identify customer links that do not appear in the augmented customer list of a firm in 2008 but do appear in 2009 and count them as link entries Links that appear in the firm's augmented customer list in both 2008 and 2009 are referred to as survivals. The table shows that the entry rate for 2008–2009 (the number of link entries in 2009 relative to the total number of links in 2008) is 10.8 %, and the exit rate during the same period is 7.4 %. Since the entry rate exceeds the exit rate, the number of links increases from 2008 to 2009 by 3.4 %. On the other hand, the survival rate for 2008-2009 (i.e., the number of surviving links between 2008 and 2009 relative to the total number of links in 2008) is 92.6 %, indicating that firms update their customer lists only partially within a year. Given that the survival rate falls to 87.2 % for the 2-year period from 2008 to 2010, 82.5 % for the 3-year period from 2008 to 2011, and 78.2 % for the 4-year period from 2008 to 2012, the survival rate for the next τ years, which we denote by $R^C(\tau)$, is estimated as

$$R^C(\tau) = 0.978 \exp(-0.056\tau). \tag{2.8}$$

Table 2.2 Turnover of customer and supplier links

Customer links	Number of links in the initial year	Net increase	Entries	Survivals	Exits
Between 2008 and 2009	867,612	29,583 (0.034)	93,540 (0.108)	803,655 (0.926)	63,957 (0.074)
Between 2008 and 2010	829,014	52,511 (0.063)	158,564 (0.191)	722,961 (0.872)	106,053 (0.128)
Between 2008 and 2011	801,508	70,835 (0.088)	210,754 (0.263)	661,589 (0.825)	139,919 (0.175)
Between 2008 and 2012	781,578	78,281 (0.100)	248,723 (0.318)	611,136 (0.782)	170,442 (0.218)
Supplier links	**Number of links in the initial year**	**Net increase**	**Entries**	**Survivals**	**Exits**
Between 2008 and 2009	864,814	19,416 (0.022)	77,154 (0.089)	807,076 (0.933)	57,738 (0.067)
Between 2008 and 2010	830,486	32,667 (0.039)	128,810 (0.155)	734,343 (0.884)	96,143 (0.116)
Between 2008 and 2011	801,210	46,418 (0.058)	173,515 (0.217)	674,113 (0.841)	127,097 (0.159)
Between 2008 and 2012	779,470	56,970 (0.073)	210,670 (0.270)	625,770 (0.803)	153,700 (0.197)

The figures in parentheses show the ratio to the number of links in the initial year

Given the above relationship, simple calculation indicates that about 45 % of links disappear over a decade and 70 % over two decades. For supplier links, the entry and exit rates for 2008–2009 are 8.9 and 6.7 %, respectively, and the survival rate is 93.3 %, indicating a slightly lower turnover than for customer links. The survival rate for the next τ years, $R^S(\tau)$, is given by

$$R^S(\tau) = 0.979 \exp(-0.050\tau). \tag{2.9}$$

Next, we examine changes in the total number of links, i.e., N^C and N^S, over time. We saw in Fig. 2.2 that the distribution of the total number of links, for both customer and supplier linkages, does not change much over the 5 years. However, this does not necessarily imply that the number of links for each firm does not change much. For example, suppose a firm increases its links from 20 in year t to 100 in year $t+1$, and another firm reduces its links from 100 to 20. In this case, the link distribution does not change at all between years t and $t+1$. To see whether underneath the stable distribution there are changes in firm links that more or less offset each other, Fig. 2.5 presents scatter plots for customer and supplier linkages respectively, showing the number of a firm's links in year t on the horizontal axis and the number of its links in year $t+1$ on the vertical axis. We see that the dots are concentrated on the 45° line

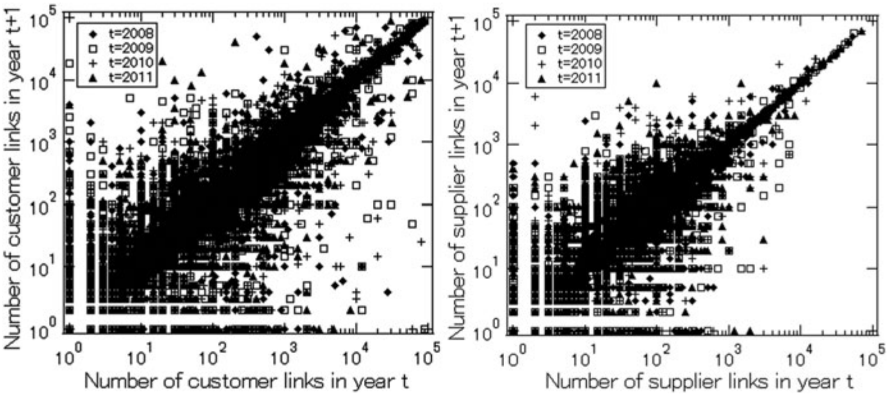


Fig. 2.5 The number of links in year *t* on the horizontal axis versus the number of links in year *t+1* on the vertical axis. The upper and lower panels are for customer links and for supplier links respectively. The figures are produced using the *total* number of links, i.e., N^C and N^S. The number of firms used in the figures is shown in Table 2.2

for both customer and supplier links, indicating that for most firms the number of links remained unchanged. At the same time, there are also dots away from the 45° line; for example, for some firms, links increase by a factor of ten or even 100, while for others they decrease by similar factors to one-tenth or one-hundredth. Comparing the scatter plots for customer and supplier links, more dots are away from the 45° line for customer links, indicating that links with customers are more volatile than those with suppliers.

To examine in more detail how firms' number of links changes over time, we show in Fig. 2.6 the distributions of the annual growth rates for the number of customer links, $\log N_i^C(t+1)/N_i^C(t)$, and for the number of supplier links, $\log N_i^S(t+1)/N_i^S(t)$, with the growth rates on the horizontal axis and the corresponding densities on the vertical axis. Note that there are eleven distributions in total in the two panels, each corresponding to a group of firms with a certain number of links in year t. For example, the distribution labeled $10^{3.5} \leq N^C(t) < 10^{4.0}$ represents the distribution of the growth rates of the number of customer links from year t to year $t+1$ for firms with a number of customer links within the indicated range.

Figure 2.6 shows the following. First, there is a clear peak in the distribution at densities corresponding to a growth rate of zero. The ratio of firms with a zero growth rate is 93.0 % for customer links and 95.2 % for supplier links. Second, each distribution has a fat upper tail. This can be seen more clearly if we compare the distributions with the dotted line representing a normal distribution with the same mean and standard deviation as observed in the data. Interestingly, the upper tail is even fatter for distributions of customer links than supplier links, suggesting again that fierce competition among firms to acquire new customers yields winners with very high growth in customer links as well as losers with very large negative growth. Third, the distributions do not depend much on the number of links in year t. To show

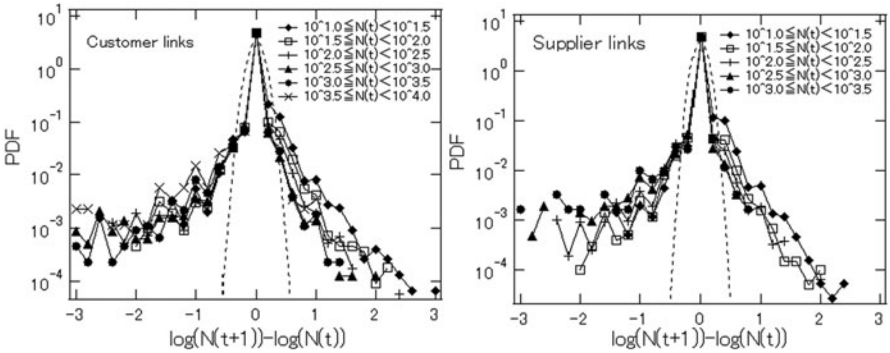
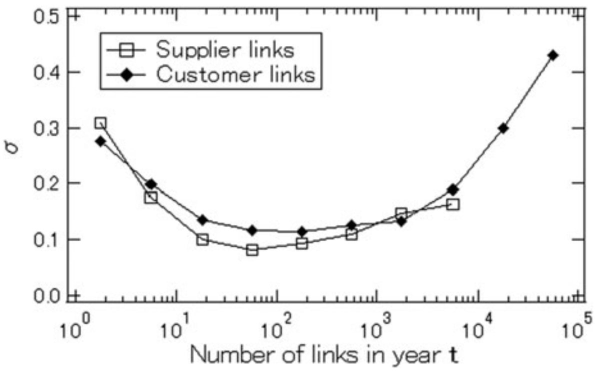


Fig. 2.6 Distributions of link growth rates from year t to year $t + 1$ for customer links (*upper* panel) and for supplier links (*lower* panel). The *dotted curve* in the *upper* panel represents a normal distribution with a standard deviation of 0.12, which is the standard deviation estimated for the growth rate of customer links, while the *dotted curve* in the *lower* panel represents a normal distribution with a standard deviation of 0.10, which is the standard deviation estimated for the growth rate of supplier links

Fig. 2.7 Relationship between the number of links in year t, which is shown on the horizontal axis, and the standard deviation of link growth rates from year t to year $t + 1$, which is shown on the vertical axis



this more clearly, Fig. 2.7 plots the number of links in year t against the standard deviation of the growth rates of links from t to $t + 1$. The figure shows that although the standard deviation is relatively high when the number of links in year t is either very small (i.e., below 10) or very large (above 10^4), it is comparatively small and almost uniform for intermediate values.

To see what this almost uniform standard deviation means, let us consider a simple Poisson type situation. We assume that the number of attempts that a firm makes to acquire new customers in $t+1$ is proportional to the number of customers the firm has in t. We denote the number of attempts by $\alpha N_i^C(t)$, where α is a positive parameter. We also assume that the probability of success for each attempt is $1/\alpha$. In this simple setting, the growth rate from t to $t + 1$ of the number of customers for a firm is, on average, unity, which is consistent with the empirical result shown in Fig. 2.6. However, the standard deviation of the link growth rates is $\sqrt{N_i^C(t)\,(1 - 1/\alpha)}$, indicating that the standard deviation is not invariant but decreases with $N_i^C(t)$ due to the law of large numbers, which is clearly inconsistent with the empirical result presented in Fig. 2.7. The result shown in Fig. 2.7 suggests that the law of large numbers does *not* hold in the data, so that the risk of losing many customers from t to $t + 1$ is not small even for firms with a large customer base in period t.

2.6 Implications for Firm Sales and Growth

2.6.1 The Relationship Between Customer Links and Firm Sales

The sales of a firm in a particular year can be decomposed into two parts: sales to other firms as intermediate output ("intermediate demand" in the terminology of input-output analysis) and sales to non-firm sectors, including consumers, the public sector, and foreign buyers ("final demand"). The intermediate demand component of a firm's sales can be further decomposed into two determinants: the number of customer links and the average size of customer links (in terms of sales). In the terminology of economics, the number of customer links is the extensive margin, while the average size of customer links is the intensive margin. An important question to be asked is which of the two margins accounts for differences in the intermediate demand component of firm sales. In the context of international trade, this issue has been addressed by a number of studies including Riccaboni and Stefano Schiavo (2010), Chaney (2008), some of which show the relative importance of the extensive margin (Chaney 2008). In the context of firm dynamics, some studies argue that the number of customer links plays a dominant role in explaining differences in firm sales (Saito et al. 2007), while certain anecdotal evidence suggests that having links of a larger size, which may reflect closer and longer-lasting ties with a particular partner, makes it possible for firms to achieve higher sales.

However, to the best of our knowledge, researchers have access neither to information that makes it possible to decompose firm sales into final and intermediate demand nor to information on the size of customer links. Our dataset does not contain that kind of information either. However, we are still able to investigate how the number of customers for a firm is related to the sales of the firm. To this end, Fig. 2.8 shows the relationship between the two, depicting the number of a firm's customers on the horizontal axis and the firm's sales on the vertical axis. More specifically, we define 14 bins of the same size in logarithm for the number of customer links and show various percentiles of the sales distribution for firms belonging to each bin, namely the 25th (×), 50th (○), 90th (▲), 99th (□), and 99.9th (◆) percentiles. As can be clearly seen in the figure, sales are positively correlated with the number of customer links. Moreover, a simple regression indicates that the median of the sales distribution in logarithm, denoted by m, depends on the number of customer links. Specifically, the relationship can be expressed as follows:

$$m \propto \frac{1}{2} \ln N^C \tag{2.10}$$

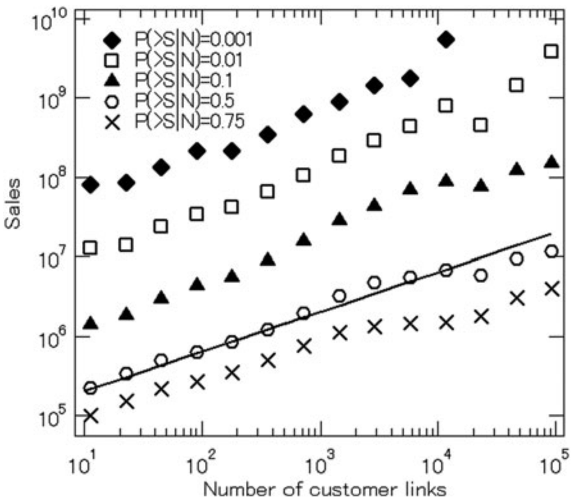


Fig. 2.8 Firm sales conditional on the number of customer links. We define 14 bins of the same size in logarithm for the number of customer links and show various percentiles of the sales distribution for firms belonging to each bin, namely the 25th (×), 50th (○), 90th (▲), 99th (□), and 99.9th (♦) percentiles. The solid straight line is a reference line with a slope of 1/2

Note that a similar linear relationship holds for the other percentiles, especially for the upper tail part, which is consistent with the results reported in the previous studies including Saito et al. (2007), Watanabe et al. (2013). Equation (2.10) implies that the variance in the log of sales is related to about 25 % of the variance in the log of the number of customer links, suggesting that the extensive margin is relatively important. At the same time, however, Eq. (2.10) also indicates that a 10 % increase in the number of customer links raises firms' sales only by 5 %, implying that other determinants of firms' sales that are not controlled for in the regression may be inversely correlated with the number of customer links. For example, the size of customer links may be negatively correlated with the number of customer links; that is, firms with a larger customer base may have customer links of a smaller size. Alternatively, firms with a larger customer base may be located more upstream in customer-suppliers chains, so that they may have less opportunity to sell their products to consumers, etc., as final output.

2.6.2 *Can Customer–Supplier Links Predict Firm Growth Correlations?*

Close interconnectedness among firms implies that an idiosyncratic shock to a firm could diffuse widely to other downstream firms through customer–supplier chains and, ultimately, result in fluctuations in the economy as a whole. As clearly demonstrated by the recent earthquake in Japan, the production activities of firms are closely correlated when these firms are "neighbors" in a customer–supplier network.

To investigate such correlation in production activities in more detail, we compute the correlation in annual sales growth between two firms, firms i and j, which

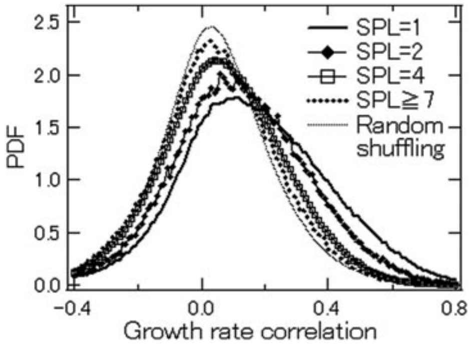


Fig. 2.9 Distributions of growth rate correlations between two firms with different shortest path lengths. The thin dotted line labeled "random shuffling" represents the distribution for the case of random shuffling in which (1) we randomly pick two years for a particular firm, swap the growth rates for the 2 years, and repeat this for other pairs of years; (2) we conduct the same random shuffling for other firms until we have completely eliminated any correlation between the growth rates for any pair of firms

is represented by ρ_{ij}. We do so for every year for all firms on the augmented customer/suppliers lists in 2008–2012 whose sales data are available for 1980–2009. (The number of firms that meet these criteria is 134,067.) We then examine how ρ_{ij} is related to the shortest path length between firms i and j. The results are shown in Fig. 2.9, which depicts the distribution of ρ_{ij} for firms that are connected by one path length (labeled "$SPL = 1$"), by two path lengths ($SPL = 2$), by four path lengths ($SPL = 4$), and by seven or more path lengths ($SPL \geq 7$). We find that ρ_{ij} is distributed around zero in the case of $SPL \geq 7$. In fact, the distribution in this case is almost identical to the distribution obtained by eliminating any correlations between firm growth rates by random shuffling,[5] which is shown by the thin dotted line, indicating that there is no statistically significant correlation between the growth rates for firms i and j. However, the distribution of ρ_{ij} moves to the right in the case of $SPL = 4$, more to the right in the case of $SPL = 2$, and even more to the right in the case of $SPL = 1$. These results indicate that there is a positive and statistically significant correlation between the growth rates for firms i and j if they are close to each other in a customer–supplier network. Simple regression shows that the growth rate correlation between firms i and j is related to the shortest path length between them as follows:

$$\langle \rho_{ij} \mid l_{ij} = l \rangle = 0.21 \exp(-0.48l) + 0.045 \tag{2.11}$$

where l_{ij} is the shortest path length between firms i and j, and $\langle \rho_{ij} \mid l_{ij} = l \rangle$ is the average of ρ_{ij} conditional on that the shortest path length between them is l. The first

[5] We eliminate growth rate correlations among firms as follows. For a particular firm, we randomly pick 2 years, swap the growth rates for the 2 years, and repeat this for other pairs of years. We do the same for all other firms until we have completely eliminated any correlation between the growth rates for any pair of firms.

term in Eq. (2.11), $\exp(-0.48l)$, indicates that the growth rate correlation decreases with l.

The positive constant term in Eq. (2.11), 0.045, indicates that the growth rates of firms i and j are positively correlated even when l is very large, implying that part of the growth rate correlations may be due to factors that have nothing to do with customer–supplier chains. In fact, the growth rate correlation for pairs of firms which are not connected at all in the network (i.e., $SPL = \infty$) is, on average, 0.056, which is close to the constant term in Eq. (2.11). To examine the relationship between the growth correlation and the shortest path length in more details, we follow the recent literature on supply chains (Foerster 2011; Watanabe et al. 2013) and assume that

$$g_t = \tilde{A} g_t + \epsilon_t \tag{2.12}$$

where g_t is a vector of firm growth rates g_{it}, where i stands for the firm, so that $g_t = (g_{1t}, g_{2t}, \ldots, g_{Nt})'$), and ϵ_t is a column vector representing shocks not stemming from customer–supplier chains ($\epsilon_t = (\epsilon_{1t}, g_{2t}, \ldots, \epsilon_{Nt})'$). $\tilde{A}$ is an $N \times N$ input-output matrix used in Eq. (2.3) with typical element $\tilde{a}_{ij}$ equal to $1/\tilde{N}_i^S$ if firm j is the supplier of firm i and zero otherwise. We assume further that ϵ_{it} can be decomposed into shocks common to all firms, such as changes in monetary and fiscal policies, and idiosyncratic shocks:

$$\epsilon_{it} = u_t + v_{it} \tag{2.13}$$

where u_t and v_{it} represent common and idiosyncratic shocks respectively, and v_{it} and v_{jt} are uncorrelated. Using Eqs. (2.12) and (2.13), we decompose growth correlations into two parts: the correlation stemming from customer–supplier linkages, and the correlation due to common shocks. We first use $\epsilon_t = (I - \tilde{A})g_t$ to recover ϵ_t. Then, we eliminate the simultaneous pairwise correlation between ϵ_i and ϵ_j by randomly exchanging ϵ_{it} and ϵ_{is} until the correlation is removed completely. We denote the uncorrelated new disturbance vector by $\hat{\epsilon}_t$ and define the new growth rate vector $\hat{g}_t$ as $\hat{g}_t = (I - \tilde{A})^{-1}\hat{\epsilon}_t$. Note that in $\hat{g}_t$ the growth rates for i and j cannot be correlated through common shocks but may be correlated through customer–supplier linkages.

The result of this exercise is presented in Fig. 2.10, where the horizontal axis shows the shortest path length, while the vertical axis depicts the growth correlation conditional on the shortest path length. The result using actual growth rate data, g_t, is represented by ♦ and shows that $\langle \rho_{ij} \mid l_{ij} = l \rangle$ decreases with l, as we saw in Eq. (2.11). Next, the result for the growth rate correlations only through linkages, which are calculated using $\hat{g}_t$, are shown by ⋄. The result indicates that $\langle \rho_{ij} \mid l_{ij} = l \rangle$ again falls with l, but this time it falls very close to zero when $l \geq 7$. Finally, we add the estimate for the growth rate correlations through common shocks, 0.045 in Eq. (2.11), to the growth rate correlations through linkages. Doing so shows that the sum of the two, which is represented by □, successfully generates the growth rate correlations observed in the data.

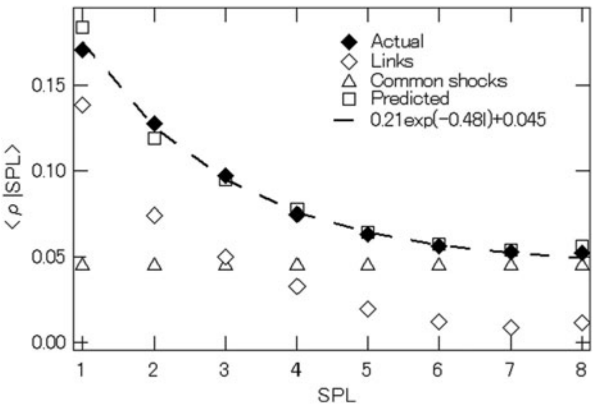


Fig. 2.10 Average of the growth rate correlations between pairs of firms conditional on the shortest path length between the pairs. The figure shows the correlations obtained from the data (♦), the correlations through common shocks (△), the correlations through customer–supplier links (◇), and the correlations through the sum of the two (□)

2.7 Conclusion

In this study, we investigated the structure and evolution of customer–supplier networks in Japan using a unique dataset that contains information on customer and supplier links for more than 500,000 incorporated non-financial firms for the 5 years starting from 2008 to 2012. Our main findings can be summarized as follows. First, we show that the number of customer links is unequal across firms in the sense that the customer link distribution is substantially skewed to the right. The upper tail of the customer distribution is much thicker than that of a lognormal distribution and close to that of a power-law distribution with an exponent of unity (i.e., it follows Zipf's law). We interpret this as implying that competition among firms to acquire new customers yields winners with a large number of customers, as well as losers with fewer customers. We also show that the distribution for the number of suppliers across firms has a power-law tail, but the associated exponent is greater than unity, indicating that the number of supplier links is less unequal than the number of customer links. Second, we find that firms are closely interconnected with each other. Specifically, the shortest path length for any pair of firms is, on average, 4.3 links. Third, we show that in our observation period the frequency of link switching is limited and that, consequently, customer–supplier networks are fairly stable over time. Our estimates indicate that the survival rate for customer links (i.e., the rate at which existing customer links survive one more year) is 92 %, while that for supplier links is 93 %. Fourth, we find that the growth rates of a pair of firms tend to be more highly correlated when the two firms are closer to each other in a customer–supplier network (i.e., the shortest path length between the two firms is shorter), suggesting that a non-negligible portion of fluctuations in firm growth stems from the propagation of microeconomic shocks—that is, shocks affecting only a particular firm—through customer–supplier chains.

In this chapter, we have focused on interfirm connections through customer–supplier relationships, but some previous studies, including Plerou et al. (2002), Bonanno et al. (2004), Ibuki et al. (2013), investigate interfirm connections through the comovement of stock prices. Specifically, these studies use stock price correlations between firms to construct an interfrim network, finding that firms with customer–supplier relationships, such as an automobile producer and a tire manufacturer, tends to be close to each other (i.e., path length is short) even in the network constructed based on stock price correlations. However, these two interfirm networks are not necessarily identical. For example, it is known that, during stock price bubbles, stock price correlations do not necessarily have one-to-one correspondence with customer–suppliers relationships. There remains much to be done regarding how these two interfirm networks are related with each other.

Acknowledgments We would like to thank Vasco Carvalho, Hiro Ishise, Makoto Nirei, Tack Yun, and partcipants at European Conference on Complex Systems held in Vienna on September 12–16, 2011 for helpful comments and suggestions, and Masahiro Miyatani for his careful explanation of the TDB data. This research forms part of a project on "Designing Interfirm Networks to Achieve Sustainable Economic Growth" carried out at Hitotsubashi University with financial support from the Ministry of Education, Culture, Sports, Science, and Technology. The dataset we use in this paper is compiled jointly by Teikoku Databank, Ltd. and the HIT-TDB project of Hitotsubashi University.

References

Acemoglu, D., Carvalho, V. M, Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica, 80*(5), 1977–2016.

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2013a). Systemic risk and stability in financial networks. No. w18727. National Bureau of Economic Research.

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2013b). The network origins of large economic downturns. No. w19230. National Bureau of Economic Research.

Atalay, E., Hortacsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences, 108*(13), 5199–5202.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science, 293*(5536), 1818–1820.

Battiston, S., Delli Gatti, D., Gallegati, M., Greenwald, B., & Stiglitz, J. E. (2007). Credit chains and bankruptcy propagation in production networks. *Journal of Economic Dynamics and Control, 31*(6), 2061–2084.

Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vanndewalle, N., & Mantegna, R. N. (2004). Networks of equities in financial markets. *The European Physical Journal B, 38,* 363–371.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks, 33*(1–7), 107–117.

Carvalho, V. M. (October 2010). *Aggregate fluctuations and the network structure of intersectoral trade*. CREI and University of Pompeu Fabra.

Carvalho, V. M. (December 2012). *A survey paper on recent developments of input–output analysis*, Report for the European Commission under the CRISIS Consortium Agreement.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review, 98*(4), 1707–1721.

Di Giovanni, J., & Levchenko, A. A. (2010). Putting the parts together: Trade, vertical linkages, and business cycle comovement. *American Economic Journal: Macroeconomics, 2*(2), 95–124.

Dupor, B. (1999). Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics, 43*(2), 391–409.

Foerster, A. T., Sarte, P. D. G., & Watson, M. W. (2011). Sectoral versus aggregate shocks: A structural analysis of industrial production. *Journal of Political Economy, 119,* 1–38.

Fujiwara, Y., & Aoyama, H. (2010). Large-scale structure of a nation-wide production network. *The European Physical Journal B, 77,* 565–580.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica, 79*(3), 733–772.

Garlaschelli, D., & Loffredo, M. I. (2004). Fitness-dependent topological properties of the world trade web. *Physical Review Letters, 93*(18), 188701.

Garlaschelli, D., & Loffredo, M. I. (2005). Structure and evolution of the world trade network. *Physica A: Statistical Mechanics and its Applications, 355*(1), 138–144.

Garlaschelli, D., Battiston, S., Castri, M., Servedio, V. D. P., & Caldarelli, G. (2005). The scale-free topology of market investments. *Physica A, 350,* 491–499.

Gourio, F., & Rudanko, L. (July 2011). Customer capital. National Bureau of Economic Research Working Paper, No. 17191.

Goyal, S. (2012). *Connections: An introduction to the economics of networks*. Princeton University Press.

Ibuki, T., Suzuki, S., & Inoue, J.-I. (2013). Cluster analysis and Gaussian mixture estimation of correlated time-series by means of multi-dimensional scaling. In F. Abergel et al. (Eds.), *Econophysics of systemic risk and network dynamics* (pp. 239-259). New Economic Windows, Springer.

Jackson, M. O. (2010). *Social and economic networks*. Princeton University Press.

Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (February 2013). Firm volatility in granular networks. Chicago Booth Research Paper 12-56.

Leontief, W. (1936). Quantitative input and output relations in the economic system of the United States. *Review of Economics and Statistics, 18*(3), 105–125.

Luttmer, E. G. J. (October 2006). Consumer search and firm growth. Federal Reserve Bank of Minneapolis Working Paper, No. 645.

Markose, S., Giansante, S., Shaghaghi, A. R. (2012). 'Too interconnected to fail' financial network of US CDS market: Topological fragility and systemic risk. *Journal of Economic Behavior and Organization, 83*(3), 627–646.

Mizuno, T., Souma, W., & Watanabe, T. (2014). The structure and evolution of buyer–supplier networks. *PLoS ONE, 9*(7): e100712. doi:10.1371/journal.pone.0100712.

Ohnishi, T., Takayasu, H., & Takayasu, M. (2010). Network motifs in inter-firm network. *Journal of Economic Interaction and Coordination, 5*, 171–180.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T., & Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data. *Physical Review E, 65,* 066126.

Riccaboni, M., & Schiavo, S. (2010). Structure and growth of weighted networks. *New Journal of Physics, 12*(2), 023003.

Saito, Y. U, Watanabe, T., & Iwamura, M. (2007). Do larger firms have more interfirm relationships? *Physica A, 383,* 158–163.

Souma, W., Fujiwara, Y., & Aoyama, H. (2003). Complex networks and economics. *Physica A, 324,* 396–401.

Stella, A. (September 2013). Firm dynamics and the origins of aggregate fluctuations. Available at SSRN 2222746.

Takayasu, M., Sameshima, S., Ohnishi, T., Ikeda, Y., Takayasu, H., & Watanabe, K. (2008). Massive economics data analysis by econophysics method-the case of companies' network structure. Annual Report of the Earth Simulator Center, April 2007–March 2008, 263–268.

Watanabe, H., Takayasu, H., & Takayasu, M. (2012). Biased diffusion on the Japanese inter-firm trading network: Estimation of sales from the network structure. *New Journal of Physics, 14*(4), 043034.

Watanabe, H., Takayasu, H., & Takayasu, M. (2013). Relations between allometric scalings and fluctuations in complex systems: The case of Japanese firms. *Physica A, 392,* 741–756.

Watts, D. J, & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393,* 440–442.

Takayuki Mizuno is associate professor at National Institute of Informatics. He also concurrently serves as associate professor at the Graduate University for Advanced Studies (SOKENDAI), as PRESTO researcher at Japan Science and Technology Agency (JST), and as senior research fellow at the Canon Institute for Global Studies. He was lecturer at Hitotsubashi University (2007–2010), associate professor at University of Tsukuba (2011–2013). His main fields of research are computational social science for early detection method of financial bubble using big data of company's financial statements and high frequency transactions of financial markets. He has published papers in PLOS ONE, Physica A, the European Physical Journal B, Physical Review E, Journal of the Physical Society of Japan, among others. He received a Ph.D. in Physics from Chuo University.

Wataru Souma is associate professor of physics at College of Science and Technology, Nihon University (2009–). He received his Ph.D. in physics from Kanazawa University in 1996 and studied High Energy Physics at Kyoto University (1998–2001). He served as researcher at Advanced Telecommunication Research Laboratories (ATR) (2001–2006), senior researcher at ATR (2006–2009), research specialist at National Institute of Information and Communications Technology (NiCT) (2006–2009), and visiting researcher at Institute of Economic Research at Hitotsubashi University (2009–2011). His main fields of research are network science (business network, citation network, innovation network, multiple network, etc.), econophysics (personal income distribution, firm size distribution, application of random matrix theory to analysis of financial time series, etc.), functional renormalization group, and quantum gravity. He is also interested in science for science policy, and is developing search engine for innovation management.

Tsutomu Watanabe has been a professor of economics at the Graduate School of Economics, The University of Tokyo, since October 2011. He has also been a professor at Hitotsubashi University (1999–2011) and a senior economist at the Bank of Japan (1982–1999). He has held visiting positions at various universities, including Kyoto University, Bocconi University, and Columbia University. He received his Ph.D. in economics from Harvard University in 1992 and did his undergraduate work at The University of Tokyo. Watanabe's main research area is monetary policy and inflation dynamics. He is known for his series of papers on monetary policy when nominal interest rates are bounded at zero; in particular, his paper on the optimal monetary policy at the zero lower bound, whose first draft appeared in 2001 and final version was published in 2005, has been widely recognized as the first paper to provide a simple description of the liquidity trap and characterize the optimal policy response to it in a setting of dynamic stochastic general equilibrium. He is an author of many books and more than 40 academic journal articles on monetary policy and international finance.

Chapter 3
Community Structure of a Large-Scale Production Network in Japan

Takashi Iino and Hiroshi Iyetomi

Abstract This chapter analyzes nationwide supplier-buyer relationship data for nearly a million firms and 4 million transactions in Japan. The production network constructed by firms through their transaction relations reflects the characteristics of economic activities in Japan. For an intuitive understanding of the network structure, we first visualize the network in three-dimensional space using a spring-electrostatic model. In this model, we replace nodes (firms) and links (transaction relations) by particles with identical charges and springs. This visualization shows that the network is highly heterogeneous, with some firms being tightly connected and forming groups, between which there are much looser connections. Such industrial communities are identified here using algorithms that maximize modularity, which measures the share of links encircled by a given partition of nodes, with reference to the expected share of intra-links for corresponding random networks with the same node partitions. Since major communities thereby detected are still very heterogeneous, the detection of communities is repeated within them. The 10 largest communities and their principal sub-communities are then characterized by areal and industry sectoral attributes of firms. In addition, how closely the sub-communities are related to each other is quantified by introducing a metric of "distance" between them. Finally, the hierarchical relationship between the communities is clarified by considering directional features of the transactions.

Keywords Production network · Community · Modularity · Visualization · Dendrogram

H. Iyetomi (✉)
Department of Mathematics, Niigata University, 8050, Ikarashi 2-no-cho, Nishi-ku, Niigata, Japan
e-mail: hiyetomi@sc.niigata-u.ac.jp

T. Iino
Faculty of Science, Niigata University, 8050, Ikarashi 2-no-cho, Nishi-ku, Niigata, Japan
e-mail: iino@ad.sc.niigata-u.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_3

3.1 Introduction

It is no exaggeration to say that the economy depends entirely on interactions between individuals. For instance, firms are connected to each other directly or indirectly through their business transactions. A firm obtains materials from suppliers (upstream firms) and sells its products to customers (downstream firms). These transactions are so essential to firms that one cannot isolate the dynamics of individual firms from the entire economic system. Firms' production activities thus give rise to a complex network, and examining economic phenomena from the perspective of networks can provide a variety of new insights into economic phenomena.

The study of complex networks is an area attracting growing research interest and is largely driven by the empirical study of real-world networks such as those in biology, computer science, economics, and sociology. A number of studies on complex networks have already been carried out from a physics standpoint (Albert and Barabási 2002; Newman 2003). These studies encompass the visualization of networks based on a physical model, the development of statistical mechanics methods for quantifying network structure, and the construction of theoretical models of network formation.

It is difficult to analyze all connections in complex networks, because of their multiplicity and complexity. Visualization, which is gaining popularity owing to the recent development of graphics technology, is a useful tool with which to illuminate structural properties of networks. Appropriate depiction of a complex network greatly helps in grasping its intricate structures by providing an intuitive understanding. That being said, visualization of course is not always an effective tool in all circumstances.

In general, real complex networks have a nonuniform link distribution. A typical example is the World Wide Web, which is not a random network at all, because websites tend to link to other sites of a similar nature. In social networks, people tend to establish a friendship with someone who is a friend of a friend rather than with a complete stranger. Such nonuniform networks are often characterized by communities, which are groups of tightly connected nodes, with links between the communities being sparse. Analysis of such communities facilitates a coarse-grained perspective on the structure of large complex networks.

To detect community structures of complex networks using algorithms, Newman proposed a measure called modularity, which is used to evaluate the performance of network division through the creation of modules (Newman 2004, 2006; Newman and Girvan 2004). The idea underlying the use of modularity is to quantify statistically unforeseen arrangements of links. The optimal decomposition of communities in a network is determined by finding the division with the largest modularity value. Various methods for optimizing modularity have been proposed (Fortunato 2010); detection of communities is a computationally demanding task, especially for large networks.

In a series of papers (Iino and Iyetomi 2011, 2012a, b; Iino et al. 2010; Kamehama et al. 2010), we studied community structure in a Japanese production network, constructed with data on almost a million firms and 4 million transaction relations. In this network, nodes and links correspond to firms and transaction relations, respectively. The scale of the network covered in our dataset is so large that it comprises the majority of all firms in Japan and hence accounts for almost all production activities. The objective of this chapter is to present results that we have already reported piecemeal in a more coherent fashion.

We have previously conducted community analysis and visualization of the entire network (Iino et al. 2010; Kamehama et al. 2010). The network was separated into tens of major communities and many minor peripheral communities. The modularity optimization therefore worked well to demonstrate that the network is strongly nonuniform. A similar study on a partial network of manufacturers based on the same data was reported in Fujiwara and Aoyama (2010). The communities extracted in the above studies often contain nodes that have common features or behaviors in networks. Social organizations tend to generate hierarchical structures. For instance, Japan has several regions, each composed of several prefectures. These communities may be characterized by lower-level social organizations. Other causes of the formation of such communities include the development of supply chains in industry sectors and the establishment of business affiliations. Business relationships between firms may also reflect how close the firms are financially, politically, and historically.

It is well known that modularity optimization often fails to identify small but important communities embedded within large communities (Fortunato and Barthélemy 2007). This indicates that dominant communities in the transaction network may have more detailed nonuniform structures. To address this resolution-limit issue of modularity, we sorted through community structures within the dominant communities of the transaction network by repeating the community detection procedure (Iino and Iyetomi 2011, 2012a). Hereafter, "community" refers to a community in the entire network, and "sub-community" to a community within that community. Furthermore, we evaluated the strength of business links between the (sub-)communities by merging every pair of those modules and calculating the resultant change in modularity, which quantifies to what extent they are coupled to each other. This evaluation enables us to consider the relative positions of the (sub-)communities in the entire (partial) network(s).

Finally, we address directional bias regarding links representing transaction flows between suppliers and buyers (Iino and Iyetomi 2012b). We compute polarization and the total number of interconnecting links for every pair of communities. We also devise a visualization method based on a spring–electrostatic model for directed networks by adding a "magnetic" interaction with an external field; greatly polarized links tend to be aligned by the field. This allows us to elucidate the structure of transaction flows in the production network at the level of communities.

The remainder of the chapter is organized as follows. The next section explains the interfirm transaction data used to analyze the structure of production networks in Japan. Section 3.3 presents the configuration of the entire network and its partials optimized in three-dimensional space, together with the method of optimization. Section 3.4 provides a brief account of how communities are detected on the basis of modularity. Section 3.5 then discusses the characteristics of the 10 largest communities and their major sub-communities in the production network in light of areal and business sectoral information of constituent firms of those (sub-)communities. Section 3.6 focuses on measuring the strength of relations between those (sub-) communities. Next, Section 3.7 discusses directional bias in transactions spanning the communities. Finally, Section 3.8 summarizes the results and provides future perspectives.

3.2 Interfirm Transaction Data in Japan

The data on firms used here were collected in 2005 by Tokyo Shoko Research, Ltd., using financial statements, corporate documents, and interview-based surveys. The dataset contains information on more than 1 million firms and 4 million transactions between firms. This compares to an estimated 2.5 million firms in total according to the census conducted by the Ministry of Internal Affairs and Communications. The data cover virtually all types of economic activity in Japan. Transaction data are filtered to remove all firms that were identified as bankrupt in the database or for which important information, such as their industry classification, was missing. The resulting network is not a connected one, although its largest connected component with $N = 773,670$ nodes and $M = 3{,}192{,}582$ links encompasses more than 99 % of total nodes and links; the remaining connected components have at most eight nodes. Therefore, in the following analysis, only the largest connected component is considered.

As shown in Fig. 3.1, the connections between firms have characteristics of a scale-free network (Fujiwara and Aoyama 2010). That is, the number of links (degree) originating from a node has a power-law distribution; there are firms acting as hubs, with a number of connections at every scale. Scale-free networks are observed in various places, including the topology of webpage links, the collaborative networks of Hollywood actors, and the United States power grid (Barabási 2002; Buchanan 2002; Newman 2003; Watts 1999). Preferential attachment is a possible mechanism for the formation of such networks; the probability of linking to a node is proportional to the number of existing links that node has.

3.3 Visualization of the Production Network

Various algorithms have been developed to visualize networks (Battista et al. 1998). Here, a spring–electrostatic model is adopted in which pairs of firms with direct transaction relations are physically connected by springs, and firms in any pair repel each other through a repulsive Coulomb force (Hu 2006). The attractive force of the spring keeps intimate firms close in space. On the other hand, the repulsive Coulomb force tends to distribute firms uniformly over the available space and prevents en-tanglement of the transaction network. We then take full advantage of a molecular dynamics (MD) method (Allen and Tildesley 1987; Frenkel and Smit 2002) for an

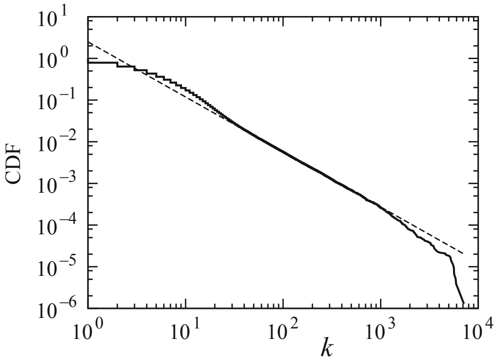


Fig. 3.1 Cumulative distribution function (CDF) of degree k, given as the sum of incoming and outgoing links for each firm in the Japanese production network. The tail of the CDF fits the power law form, $k^{-\mu}$, well, with $\mu = 1.32$ (*dashed line*)

optimized configuration of nodes in the model. The ground state in the model is a leading candidate for this configuration. The MD simulation works well to reproduce an ordered structure, with the lowest-energy forms such as crystals of materials generated through slow cooling, starting from any initial configuration. We expect that the simulation is also successful in visualizing the network. However, the MD simulation of a system consisting of charged particles on the order of a million is not a simple task, because of the long-range nature of the Coulomb force. The computational task is of the order of N^2, without any original algorithms. To speed up the computation, we implemented the hierarchical tree algorithm (Barnes and Hut 1986; Pfalzner 1996) in our MD code.

The interaction force between nodes l and m for the model is explicitly written as

$$F\left(r_{lm}\right) = -k_{lm} r_{lm} + \frac{q_l q_m}{r_{lm}^2}, \tag{3.1}$$

where k_{lm} is the spring constant for the attraction between the nodes. If the nodes are directly connected, $k_{lm} = k$; otherwise $k_{lm} = 0$. q_l denotes the Coulomb charge for node l. Here we neglect the direction of links (flow of goods or money) and assume that q_l and k_{lm} take on identical values for every node and pair, respectively. We use the velocity Verlet algorithm (Allen and Tildesley 1987; Frenkel and Smit 2002) to integrate the equation of motion of a system of classical particles interacting through Eq. 3.1.

The nodes were distributed randomly in a cubic box at the initial time. The given configuration immediately led to a high-temperature state in which all nodes moved almost freely. We then gradually decreased the temperature of the system through a reduction of the kinetic energies of nodes to reach a zero-temperature state. The relative magnitudes of the potential parameters control the final structure obtained, whereas the absolute magnitudes are completely irrelevant. We confirmed that the final result was essentially independent of the initial configuration and cooling process adopted. An example of the relaxation process is provided in Fig. 3.2.

For the analysis, firms are divided into eight broad industry categories: construction, manufacturing, services, real estate, transport, wholesale & retail trade,

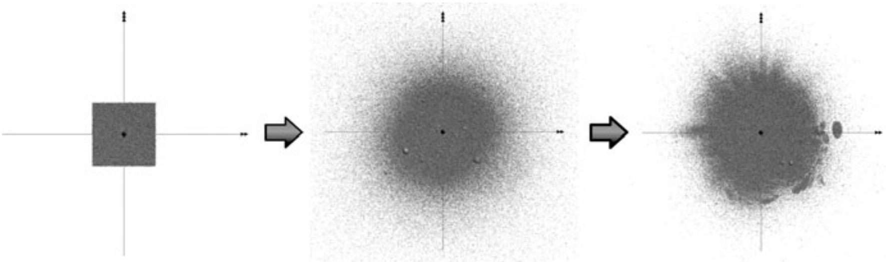


Fig. 3.2 Relaxation process in the MD simulation, where only nodes are drawn; initial state (*left*), intermediate state (*middle*), and final state with zero temperature (*right*). (Adapted from Kamehama et al. 2010)

information & communication, and other. Among these, construction, manufacturing and wholesale & retail trade are the largest sectors, comprising almost 80 % of all firms. Figure 3.3 shows how firms in each sector are distributed in the network and provides an idea of the relative positions of the sectors in the economic system. Roughly speaking, the sectors are placed as follows. Firms in wholesale & retail trade gather on the front side of the space in Fig. 3.3, while most firms in manufacturing are positioned in the vicinity of the center, and firms in construction occupy the back side. Firms belonging to services and real estate are spread over the entire space. In contrast, firms in transport and information & communication form a core at the center. The figure is thus a kind of industrial map of Japan.

It can also be seen that each sector has its own structural characteristics. For instance, construction has a brush-like structure. This indicates that a hierarchical relationship (contractors, sub-contractors, sub-sub contractors, and so on) prevails in the sector. In contrast, firms in manufacturing are tightly connected to each other, forming a dense cluster. Wholesale & retail trade is characterized by a number of mushroom-like structures of various sizes. This shows that wholesalers tend to have relationships with many associated retailers. Typical structures of the three sectors are illustrated in Fig. 3.4.

The three structures corresponding to the major sectors are individually depicted in Fig. 3.5. These network structures were optimized in separate MD runs. The characteristic properties (recall Fig. 3.4) of the three sectors are depicted more clearly here. The partial network of the manufacturing sector was also visualized by Fujiwara and Aoyama (2010).

3.4 Modularity Maximization

To find community structures in a network, we need a criterion with which to evaluate the density of connections of a group for a given partition. Newman proposed "modularity" as such a criterion. He extracted community structures in networks by finding the division that gives the largest modularity value.

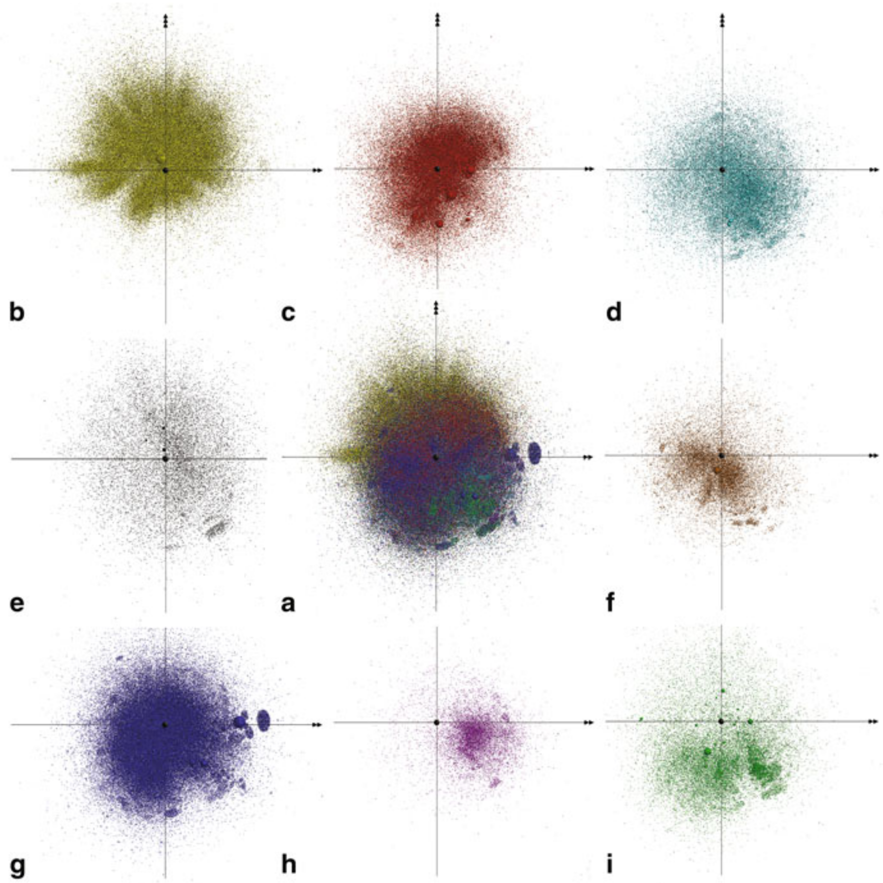


Fig. 3.3 Optimized network structure of all sectors of the economy. There are eight sectors designated by different colors; the central image (**a**) is identical to the rightmost drawing in Fig. 3.2. The sector classification is **b** construction, **c** manufacturing, **d** services, **e** real estate, **f** transport, **g** wholesale & retail trade, **h** information & communication, and **i** other. (Adapted from Kamehama et al. 2010)

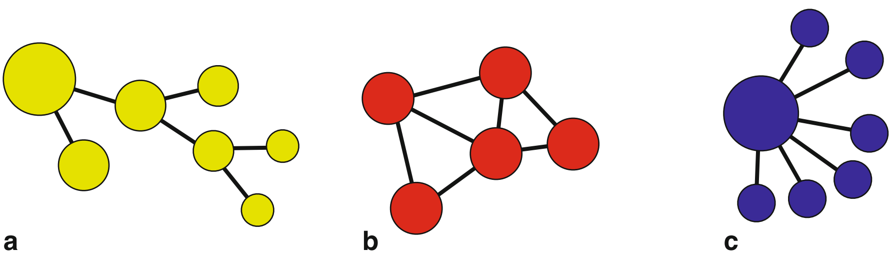


Fig. 3.4 Schematic drawing of typical transaction structures in three sectors: **a** construction, **b** manufacturing, and **c** wholesale & retail trade. (Adapted from Kamehama et al. 2010)

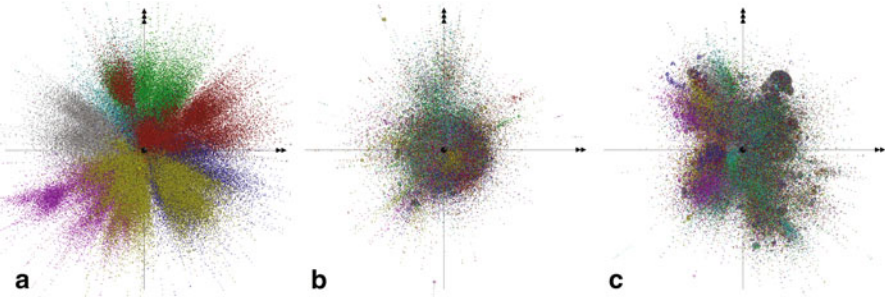


Fig. 3.5 Drawings of partial networks for three sectors: **a** construction, **b** manufacturing, and **c** wholesale & retail trade. Networks are colored to distinguish regions. (From Kamehama et al. 2010)

We begin by explaining modularity. We assume a network V is divided into L subsets $\{V_1, V_2, \cdots, V_L\}$ that do not overlap and are not empty. We define the fraction e_{ij} of the number of links that connect nodes in subset V_i and nodes in subset V_j as

$$e_{ij} = \frac{1}{2M} \sum_{l \in V_i} \sum_{m \in V_j} A_{lm} \tag{3.2}$$

The number of links in network V is represented by M. The symbol A_{lm} represents an element of the adjacency matrix for V:

$$A_{lm} = \begin{cases} 1 & \text{(if a link exists between node } l \text{ and node } m\text{)} \\ 0 & \text{(otherwise)} \end{cases} \tag{3.3}$$

We also define the fraction a_i of the number of links that connect to subset V_i from the whole part of the network as

$$a_i = \frac{1}{2M} \sum_{l \in V_i} \sum_{m \in V} A_{lm} = \sum_j e_{ij}. \tag{3.4}$$

Using e_{ij} and a_i so defined, we define modularity Q for a given partition $\{V_1, V_2, \cdots, V_L\}$ as

$$Q = \sum_{i=1}^{L} Q_i = \sum_{i=1}^{L} \left(e_{ii} - a_i^2\right). \tag{3.5}$$

In Eq. 3.5, e_{ii} represents the link density within subset V_i, and is canceled out by a_i^2 when subsets $\{V_1, V_2, \cdots, V_L\}$ are selected randomly; the term a_i^2 gives the expectation value of link density in subset V_i. Newman's modularity thus compares actual link density in a subset with its expectation value. A situation in which $Q \simeq 0$ is such that the network has no statistically significant communities compared with the randomly connected network, and $Q \simeq 1$ corresponds to a network that is almost perfectly partitioned into modules.

It is impossible to search for all the possibilities for the optimized division of a large network such as the transaction network examined here, because modularity maximization is an NP-hard problem (Brandes et al. 2008). For practical purposes, approximation methods are used, including greedy agglomeration (Blondel et al. 2008; Clauset et al. 2004; Newman 2004), simulated annealing (Guimerà and Amaral 2005; Medus et al. 2005; Reichardt and Bornholdt 2006), extremal optimization (Duch and Arenas 2005), and spectral methods (Newman 2006). We adopt a bisection method (Iino et al. 2010), which has built-in simulated annealing and requires significantly more computational time than the greedy algorithm. Although there is a tradeoff between the computational time and performance of optimization methods, the bisection method can detect communities in the sub-million transaction network in a realistic time, with a much more optimized Q value. Details of the computational method are provided in Iino et al. (2010).

3.5 Decomposed Communities and Sub-Communities

We extracted 118 communities in the transaction network with $Q_{\max} = 0.654$. The community structure thus obtained is clearly apparent in Fig. 3.6a. The community size distribution is shown in Fig. 3.7, where the size of communities is measured by how many firms they contain. There are 17 large communities with over 10, 000 nodes and about 90 tiny communities with less than 20 nodes. The top 10 communities cover about 71 % of nodes in the entire network. Table 3.1 shows the characteristics of these top 10 communities in terms of the prefecture and industry in which the firms in these communities operate. Table 3.2 provides the key for the industry abbreviations used in Table 3.1. The industry classification follows the Japan Standard Industrial Classification.

We see that each of the dominant communities is strongly determined by the regions and industry sectors of their constituent firms. For instance, the largest community to a large extent consists of manufacturers and wholesalers of machinery. The second, fourth, fifth, sixth, and ninth largest communities consist of construction groups. Among these, the fourth and ninth largest communities show particularly strong localization. Specifically, more than 90 % of firms in the fourth largest community are based in Kyushu and more than 90 % of firms in the ninth largest community are based in the Tohoku area. The third and eighth largest communities mainly consist of firms dealing with food and apparel, respectively. The seventh largest community mainly comprises transport firms. Finally, the tenth largest community consists of firms involved in printing and the manufacture of paper products.

The transaction network is very nonuniform within the dominant communities, indicating the existence of sub-communities inside them. We therefore repeat the community detection procedure for each community to elucidate this nonuniform structure in the network.

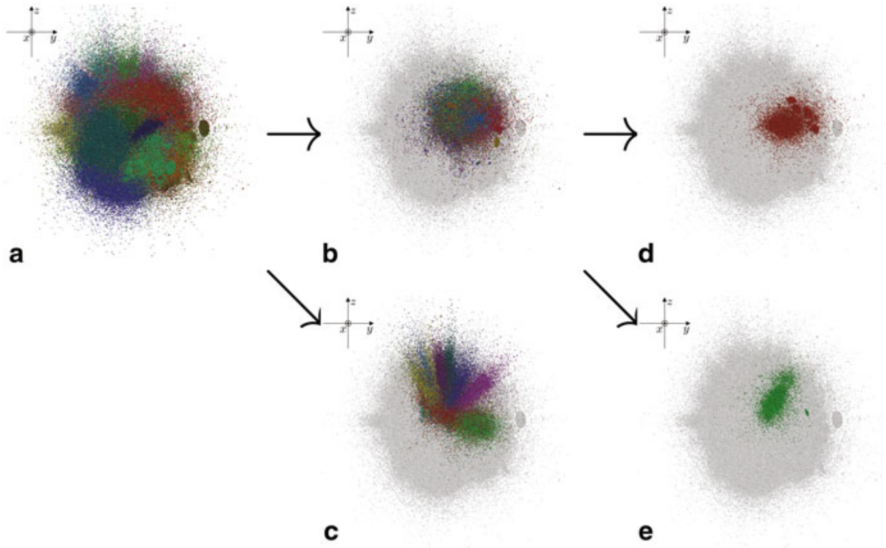
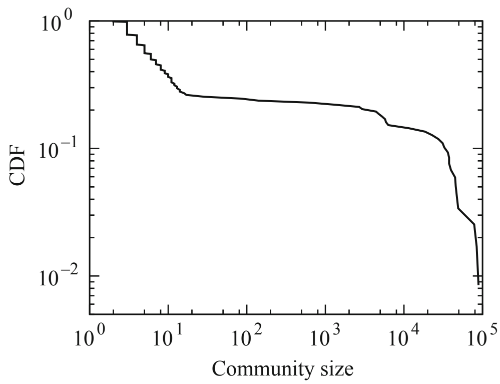


Fig. 3.6 Japanese transaction network drawn in three-dimensional space using a spring-electrostatic model. *Dots* represent nodes (firms) whose communities are distinguished by different colors. Image **a** shows the entire network, identical to the rightmost drawing in Fig. 3.2. Images **b** and **c** show only the first and second largest communities, respectively. Those communities are further decomposed into sub-communities, as shown by different colors. The first and second largest sub-communities within the largest community are selected and illuminated in (**d**) and (**e**), respectively. (From Iino and Iyetomi 2012a)

Fig. 3.7 Cumulative distribution function (CDF) for the size of communities in the transaction network, obtained employing a bisection method; the size of a community is measured in terms of the number of firms within the community. (Adapted from Iino and Iyetomi 2012a)



The results of the sub-community analysis applied to the top 10 communities are summarized in Table 3.3. The maximized values of $Q^{(i)}$ for the sub-networks are either similar to or much greater than $Q_{\max} = 0.654$, the maximized modularity for the entire network. This indicates that the major communities have considerable modular structures at the resolution level of sub-communities. In particular, half of those communities with $Q^{(i)}_{\max} > 0.7$, which are construction-based or transport-based

Table 3.1 Characteristics of firms in the 10 largest communities detected for the transaction network. The top three major prefectures and industries (see Table 3.2 for a key to the industry abbreviations) for each community are listed. Decimals in parentheses are the employment shares of firms with the specified characteristics. (From Iino and Iyetomi 2012a)

Rank	Size	Prefecture	Industry
1	88,840	Tokyo (0.189) Aichi (0.120) Osaka (0.110)	M-GM (0.144) W-ME (0.124) M-FM (0.105)
2	84,280	Niigata (0.117) Tokyo (0.094) Aichi (0.086)	C-GE (0.400) C-SP (0.228) C-EI (0.075)
3	78,529	Tokyo (0.110) Hokkaido (0.088) Aichi (0.055)	W-FB (0.262) M-FO (0.172) R-FB (0.137)
4	48,903	Fukuoka (0.298) Kagoshima (0.137) Kumamoto (0.134)	C-GE (0.359) C-SP (0.146) C-EI (0.124)
5	47,085	Aichi (0.135) Kanagawa (0.121) Tokyo (0.113)	C-GE (0.362) C-SP (0.205) W-BM (0.094)
6	45,622	Tokyo (0.152) Kanagawa (0.112) Osaka (0.098)	C-EI (0.473) C-GE (0.101) W-ME (0.072)
7	44,736	Tokyo (0.167) Aichi (0.094) Osaka (0.083)	T-RF (0.163) R-MB (0.123) R-FH (0.120)
8	39,591	Tokyo (0.180) Osaka (0.118) Aichi (0.080)	R-DA (0.251) W-TA (0.201) M-AP (0.132)
9	37,524	Fukushima (0.197) Aomori (0.167) Miyagi (0.154)	C-GE (0.357) C-SP (0.159) C-EI (0.089)
10	37,488	Tokyo (0.320) Osaka (0.096) Aichi (0.068)	M-PR (0.175) W-MI (0.127) M-PP (0.077)

groups, are expected to have clear sub-community structures from the standpoint of modularity.

To elucidate the characteristics of the sub-communities detected, we examine attributes of firms within the sub-communities, as we did for firms within the communities. The sub-communities were likewise characterized by geographic regions or industries. The first, seventh and tenth largest communities were decomposed into groups with more detailed industry classifications. The remainder had more or

Table 3.2 Key to industry abbreviations. The industry classification follows the Japan Standard Industrial Classification. Although wholesale & retail trade are merged into one category in the standard classification, we distinguish the two sectors here. (From Iino and Iyetomi 2012a)

Abbr. & Major group & Division		
C-EI	Equipment installation work	Construction
C-GE	Construction work, general, including public and private construction work	
C-SP	Construction work by specialist contractor, except equipment installation work	
M-AP	Manufacture of apparel and other finished products made from fabrics and similar materials	Manufacturing
M-BT	Manufacture of beverages, tobacco and feed	
M-EM	Manufacture of electrical machinery, equipment and supplies	
M-EP	Electronic parts and devices	
M-FM	Manufacture of fabricated metal products	
M-FO	Manufacture of food	
M-GM	manufacture of general machinery	
M-PL	Manufacture of plastic products, except otherwise classified	
M-PP	Manufacture of pulp, paper and paper products	
M-PR	Printing and allied industries	
M-RU	Manufacture of rubber products	
M-TR	Manufacture of transportation equipment	
R-DA	Retail trade (dry goods, apparel and apparel accessories)	Retail trade
R-FB	Retail trade (food and beverages)	
R-FH	Retail trade (furniture, household utensil and household appliance)	
R-MB	Retail trade (motor vehicles and bicycles)	
T-RF	Road freight transport	Transport
W-BM	Wholesale trade (building materials, minerals and metals, etc.)	Wholesale trade
W-FB	Wholesale trade (food and beverages)	
W-ME	Wholesale trade (machinery and equipment)	
W-MI	Miscellaneous wholesale trade	
W-TA	Wholesale trade (textile and apparel)	

less regional characteristics. The fourth and ninth largest communities were decomposed according to prefecture.[1] Figure 3.6b through 3.6d illustrate how the largest

[1] Japan has eight regions consisting of several neighboring prefectures, with the exception of Hokkaido, which forms its own region. The total number of prefectures is 47.

Table 3.3 Results of the sub-community analysis for the top 10 communities listed in Table 3.1. (From Iino and Iyetomi 2012a)

Rank	Number of sub-communities	Optimized modularity	Characteristics of sub-communities
1	48	0.525	Industrial
2	57	0.728	Regional
3	88	0.598	Regional
4	67	0.729	Prefectural
5	76	0.722	Marginally regional
6	116	0.685	Marginally regional
7	69	0.772	Industrial
8	49	0.638	Marginally regional
9	48	0.740	Prefectural
10	85	0.629	Industrial

and second largest communities were decomposed into sub-communities. Detailed structures are clearly evident in the network visualized by the spring–electrostatic model.

Here, we concentrate on exploring the internal structures of the three principal communities. We first focus on the region and industry of firms in the top 10 sub-communities of the largest community. Table 3.4 summarizes the results. We see that the largest community is separated into sub-communities, each of which has its own industry characteristics. The largest sub-community contains major manufacturers of electrical appliances such as Matsushita, Toshiba, and Hitachi as hub nodes. The second largest sub-community is an automobile manufacturing industry cluster, led by Toyota and its group firms such as Denso and Aishin. The third largest sub-community, a machinery-oriented group, contains Amada, Fanuc, and NSK. The fourth largest sub-community is another automobile manufacturing industry group, in which Honda, Yamaha, and Nissan are core members. The fifth largest sub-community is characterized by hub nodes that include firms such as Fujifilm, Asahi Glass, and Shimadzu. As shown in Table 3.5, sub-communities within the second largest community are defined to a large extent by the geographic regions in which they are located. This can be understood by recalling that this community is a cluster of firms in the construction industry, in which local connections are important. Table 3.6 shows that the third largest community is also separated into sub-communities characterized by well-defined geographic regions. The original community has no such regional characteristics. This community is a food industry cluster, as mentioned earlier.

Table 3.4 Characteristics of firms in the 10 largest sub-communities detected for the sub-network of the largest community. Notations are the same as those in Table 3.1. (From Iino and Iyetomi 2012a)

Rank	Size	Prefecture	Industry
1	17,020	Tokyo (0.329) Kanagawa (0.131) Osaka (0.101)	W-ME (0.157) M-EP (0.107) M-EM (0.095)
2	11,152	Aichi (0.722) Gifu (0.073) Mie (0.048)	M-GM (0.202) M-FM (0.133) W-ME (0.095)
3	9,872	Osaka (0.199) Tokyo (0.135) Shizuoka (0.081)	M-GM (0.290) W-ME (0.200) M-FM (0.118)
4	9,561	Shizuoka (0.237) Tokyo (0.157) Kanagawa (0.103)	M-GM (0.154) M-FM (0.121) M-TR (0.112)
5	7,337	Tokyo (0.249) Osaka (0.135) Kyoto (0.110)	W-ME (0.186) M-EM (0.144) M-GM (0.132)
6	6,781	Nagano (0.293) Tokyo (0.172) Kanagawa (0.067)	M-GM (0.196) M-FM (0.160) M-EM (0.071)
7	6,356	Tokyo (0.264) Osaka (0.133) Ibaraki (0.105)	M-FM (0.102) M-GM (0.096) W-ME (0.094)
8	3,773	Tokyo (0.251) Osaka (0.172) Gunma (0.101)	M-PL (0.250) M-GM (0.096) M-FM (0.082)
9	3,397	Niigata (0.367) Tokyo (0.102) Osaka (0.089)	M-FM (0.214) W-ME (0.125) W-MI (0.120)
10	2,957	Tokyo (0.214) Osaka (0.186) Hyogo (0.085)	M-RU (0.143) W-ME (0.102) M-GM (0.096)

3.6 Inter-Community and Inter-Sub-Community Relationships

The question we examine in this section is how strongly (sub-)communities are connected with each other, which allows us to delve into the nonuniform structure of the transaction network.

Table 3.5 Characteristics of firms in the 10 largest sub-communities detected for the sub-network of the second largest community. Notations are the same as those in Table 3.1. (From Iino and Iyetomi 2012a)

Rank	Size	Prefecture	Industry
1	14,135	Tokyo (0.321) Kanagawa (0.187) Osaka (0.107)	C-SP (0.349) C-GE (0.192) W-BM (0.067)
2	9,941	Niigata (0.966) Tokyo (0.010) Shizuoka (0.002)	C-GE (0.402) C-SP (0.207) C-EI (0.119)
3	7,978	Aichi (0.765) Gifu (0.160) Mie (0.012)	C-GE (0.384) C-SP (0.290) W-BM (0.056)
4	7,215	Osaka (0.311) Hyogo (0.284) Kyoto (0.160)	C-GE (0.525) C-SP (0.135) W-BM (0.067)
5	6,654	Chiba (0.236) Tokyo (0.204) Kanagawa (0.090)	C-GE (0.490) C-SP (0.162) W-BM (0.065)
6	6,633	Shizuoka (0.946) Kanagawa (0.011) Aichi (0.011)	C-GE (0.364) C-SP (0.288) C-EI (0.129)
7	5,522	Tochigi (0.615) Gunma (0.162) Ibaraki (0.133)	C-GE (0.439) C-SP (0.195) C-EI (0.062)
8	4,485	Ishikawa (0.947) Toyama (0.009) Osaka (0.007)	C-GE (0.420) C-SP (0.200) C-EI (0.122)
9	4,436	Fukui (0.955) Ishikawa (0.008) Shiga (0.007)	C-GE (0.456) C-SP (0.150) C-EI (0.131)
10	3,976	Kanagawa (0.779) Tokyo (0.126) Chiba (0.020)	C-GE (0.479) C-SP (0.181) W-BM (0.071)

3.6.1 "Distance"

We begin by introducing the modularity matrix (Newman 2006). The element B_{lm} of this matrix between nodes l and m for a network with a total number of M links

Table 3.6 Characteristics of firms in the 10 largest sub-communities detected for the sub-network of the third largest community. Notations are the same as those in Table 3.1. (From Iino and Iyetomi 2012a)

Rank	Size	Prefecture	Industry
1	10,907	Hokkaido (0.169) Tokyo (0.103) Shizuoka (0.069)	W-FB (0.387) M-FO (0.222) R-FB (0.055)
2	9,082	Tokyo (0.203) Aichi (0.128) Osaka (0.088)	M-FO (0.267) W-FB (0.250) R-FB (0.110)
3	7,574	Tokyo (0.178) Osaka (0.070) Aichi (0.067)	W-FB (0.257) R-FB (0.231) M-BT (0.128)
4	6,809	Fukuoka (0.235) Kagoshima (0.172) Miyazaki (0.141)	W-FB (0.230) M-FO (0.216) R-FB (0.120)
5	6,663	Hiroshima (0.140) Kagawa (0.126) Okayama (0.116)	W-FB (0.263) M-FO (0.204) R-FB (0.127)
6	5,294	Fukushima (0.204) Iwate (0.163) Aomori (0.163)	W-FB (0.198) M-FO (0.195) R-FB (0.155)
7	5,024	Tokyo (0.166) Kanagawa (0.084) Osaka (0.069)	W-FB (0.365) R-FB (0.181) M-FO (0.136)
8	4,947	Hokkaido (0.886) Tokyo (0.031) Osaka (0.007)	R-FB (0.276) W-FB (0.176) M-FO (0.133)
9	3,579	Tokyo (0.225) Tochigi (0.179) Gunma (0.131)	W-FB (0.299) M-FO (0.183) R-FB (0.084)
10	3,113	Aichi (0.220) Ishikawa (0.184) Toyama (0.165)	W-FB (0.275) M-FO (0.209) R-FB (0.119)

is defined as

$$B_{lm} = A_{lm} - \frac{k_l k_m}{2M}, \tag{3.6}$$

where A_{lm} is the corresponding element of the adjacency matrix of the network and $k_l = \sum_m A_{lm}$ is the degree of node l, the total number of links associated with the node. The modularity Q for a given partition of the network can be expressed in terms of the modularity matrix as

$$Q = \frac{1}{2M} \sum_{i=1}^{L} \sum_{l \in V_i} \sum_{m \in V_i} B_{lm}, \tag{3.7}$$

where we assume that all network nodes are partitioned into L groups denoted by V_i ($i = 1, 2, \cdots, L$). Maximization of Q decomposes the network into groups in an optimal way.

We project the modularity matrix onto (sub-)community space by taking the partial summation within the groups as

$$q_{ij} = \frac{1}{2M} \sum_{l \in V_i} \sum_{m \in V_j} B_{lm} = e_{ij} - a_i a_j, \tag{3.8}$$

where e_{ij} and a_i are defined by Eqs. (3.2) and (3.4), respectively. The quantity e_{ij} measures the fraction of links connecting two groups V_i and V_j, and a_i represents the fraction of links associated with V_i. If links are random in the network under the constraint that the degree of each node is fixed, the expected value of e_{ij} for such a random network is given by $a_i a_j$. We refer to the matrix defined by $\boldsymbol{q} = \{q_{ij}\}$ as a reduced modularity matrix. We note that the trace of the reduced modularity matrix amounts to the modularity Q:

$$Q = \mathrm{Tr}\boldsymbol{q} = \sum_{i=1}^{L} q_{ii}. \tag{3.9}$$

As shown below, the off-diagonal element q_{ij} ($i \neq j$) of the reduced modularity matrix is related to the increment of the modularity when two groups are combined into one, with the remaining groups untouched. We compare modularity Q of a partition $C = \{V_1, V_2, \cdots, V_L\}$ with modularity Q' of the resulting partition C', in which two groups V_i and V_j in C are merged into $V_h = V_i \cup V_j$. Link density within group V_h is calculated as

$$\begin{aligned} e_{hh} &= \frac{1}{2M} \sum_{l \in V_i \cup V_j} \sum_{m \in V_i \cup V_j} A_{lm} \\ &= \frac{1}{2M} \left(\sum_{l \in V_i} \sum_{m \in V_i} A_{lm} + \sum_{l \in V_j} \sum_{m \in V_j} A_{lm} + \sum_{l \in V_i} \sum_{m \in V_j} A_{lm} + \sum_{l \in V_j} \sum_{m \in V_i} A_{lm} \right) \\ &= e_{ii} + e_{jj} + e_{ij} + e_{ji}. \end{aligned} \tag{3.10}$$

The fraction of links connected to V_h is given by

$$a_h = \frac{1}{2M} \sum_{l \in V_i \cup V_j} k_l = \frac{1}{2M} \left(\sum_{l \in V_i} k_l + \sum_{l \in V_j} k_l \right) = a_i + a_j. \tag{3.11}$$

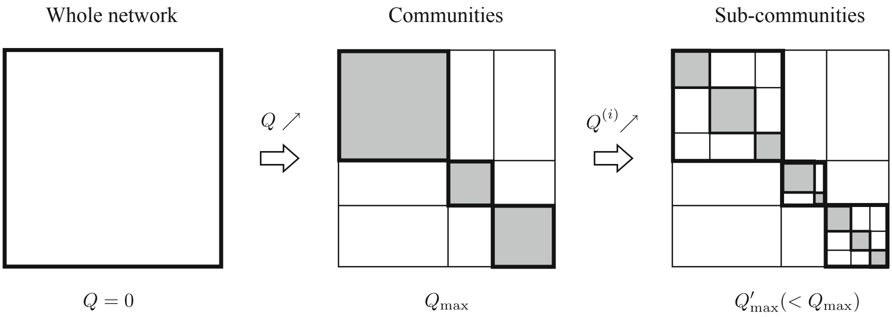


Fig. 3.8 Schematic view of the process to detect communities and sub-communities of a network in an adjacency matrix representation. The modularity Q of the entire network is first maximized to determine an optimized community structure with $Q_{\max}$. Maximization of modularity $Q^{(i)}$ for the sub-network of community i then further divides the group into components. The resulting modularity $Q'_{\max}$ for the entire network resolved to the sub-community level is less than $Q_{\max}$. (From Iino and Iyetomi 2012a)

Increment $\Delta Q = Q' - Q$ of modularity is thus calculated as

$$\begin{aligned}\Delta Q &= e_{hh} - a_h^2 - \left(e_{ii} - a_i^2 + e_{jj} - a_j^2\right) \\ &= e_{ij} + e_{ji} - a_i a_j - a_j a_i \\ &= q_{ij} + q_{ji} \\ &= 2q_{ij},\end{aligned} \tag{3.12}$$

where the last line is valid only for undirected networks, with $q_{ij} = q_{ji}$.

The reduced modularity matrix for the communities obtained by maximizing the modularity should have *negative* off-diagonal elements. Otherwise, one could obtain a larger modularity value by merging the two communities associated with the negative element into one. In contrast, the reduced modularity matrix for the sub-communities has *positive* off-diagonal elements. This statement of course depends on the definition of the reduced modularity matrix for the sub-communities. In this work, the matrix was always calculated for the entire network with all links, including those across the communities. Partitioning of nodes by the sub-communities is therefore not optimal because there is room for optimization of the modularity by unifying two of them (that is, $\Delta Q > 0$ in Eq. 3.11); see Fig. 3.8.

A large value of q_{ij} indicates that two (sub-)communities i and j are strongly related, irrespective of sign. For communities with $q_{ij} < 0$, $\Delta Q \simeq 0$. Thus, merging the two communities does not impair the optimized decomposition at all. This is in fact interpreted as a sign that they are tightly connected. For sub-communities with $q_{ij} > 0$, the merging of the two sub-communities greatly increases modularity, which can be regarded as a manifestation of their strong coupling.

Here, we define the distance d_{ij} between (sub-)communities i and j in the reduced modularity matrix as

$$d_{ij} = \max(q_{ij}) - q_{ij}, \tag{3.13}$$

Fig. 3.9 Reduced modularity matrix for the 10 largest communities in the entire network. (Adapted from Iino and Iyetomi 2012a)

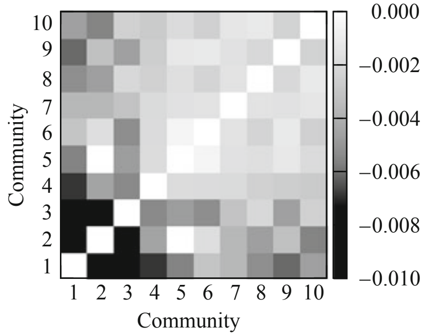
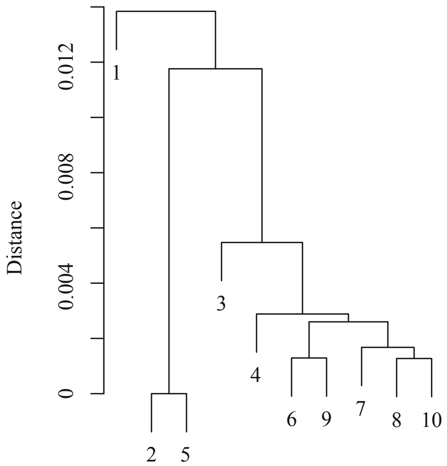


Fig. 3.10 Dendrogram derived from the reduced modularity matrix in Fig. 3.9. (Adapted from Iino and Iyetomi 2012a)



which considers the fact that distance should be inversely related to community and sub-community closeness and take a positive value. The more closely related two groups are, the shorter is their distance. This distance is extensively used to evaluate the strength of the relationship between (sub-)communities throughout the remainder of this chapter.

3.6.2 Dendrograms

We now elucidate the strength of connections between the dominant communities in Table 3.1. The reduced modularity matrix for the communities, calculated using Eq. 3.8, plays a key role. The numerical result is depicted in Fig. 3.9. To obtain a more comprehensive view of the relationship between the communities, we constructed the dendrogram shown in Fig. 3.10. This is a result of the cluster analysis with distance defined by Eq. 3.13. Each end of the branches in the dendrogram is labeled by the

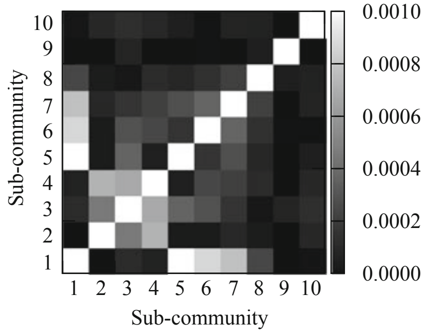


Fig. 3.11 Reduced modularity matrix for major sub-communities within the largest community. (Adapted from Iino and Iyetomi 2012a)

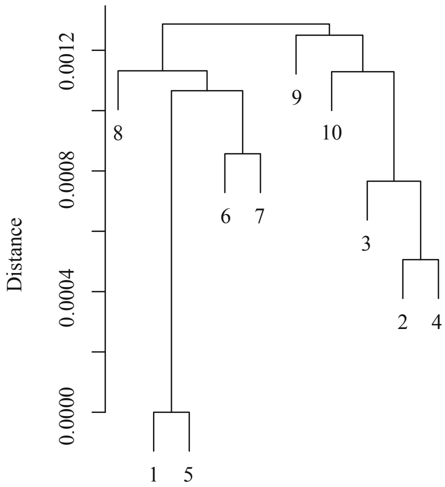


Fig. 3.12 Dendrogram corresponding to the reduced modularity matrix in Fig. 3.11. (Adapted from Iino and Iyetomi 2012a)

size rank of the corresponding community. Figure 3.10 reveals that the Japanese economy has a trilateral structure. One can separate the entire system into three groups by choosing any cutoff distance within a wide range between 0.006 and 0.011. The largest community itself forms one group. This is the manufacturing sector that leads the Japanese economy, as will be ascertained by the sub-community analysis. The tightly connected second and fifth communities form construction-based groups. The rest are assembled into the third group, in which the communities of the food industry (third), transport (seventh), apparel (eighth) and printing (10th) participate, together with the fourth, sixth, and ninth communities of construction. Therefore, the construction sector, one of the leading economic players in Japan, turns out to be separated into two groups.

The reduced modularity matrix for the sub-communities within the largest community is depicted in Fig. 3.11. The dendrogram derived from the matrix is shown in Fig. 3.12. It is widely known that the automotive and electrical machinery industries form two main flows of production in the manufacturing sector in Japan. This fact

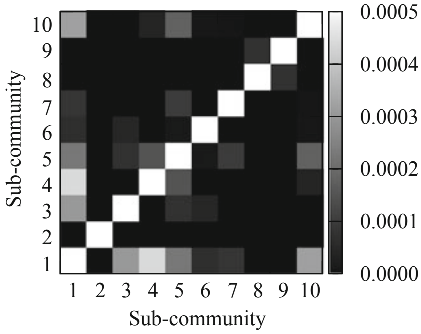


Fig. 3.13 Reduced modularity matrix for major sub-communities within the second largest community. (Adapted from Iino and Iyetomi 2012a)

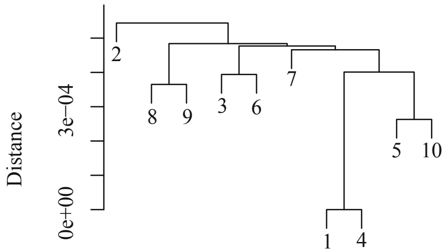


Fig. 3.14 Dendrogram corresponding to the reduced modularity matrix in Fig. 3.13. (Adapted from Iino and Iyetomi 2012a)

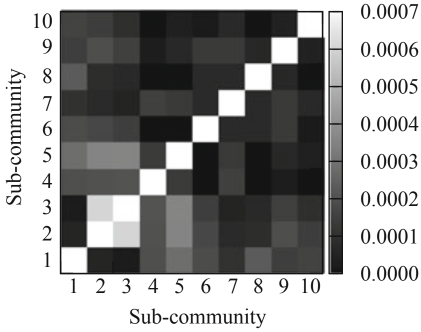


Fig. 3.15 Reduced modularity matrix for major sub-communities within the third largest community. (Adapted from Iino and Iyetomi 2012a)

is reflected well in the dendrogram with two fundamental branches: the left side features the electrical machinery industry and the right side the automotive industry.

The reduced modularity matrix and its associated dendrogram for the second largest community are given in Figs. 3.13 and 3.14 , respectively. The flat internal structure of the community, as shown in the dendrogram, strongly contrasts with the hierarchical structure of the largest community revealed in Fig. 3.12. This indicates that the sub-communities occupy relatively equal positions. This relationship between the sub-communities conforms to the radial shape of the community, as shown in Fig. 3.6c.

The reduced modularity matrix for the sub-communities within the third largest community is shown in Fig. 3.15, along with the associated dendrogram in Fig. 3.16.

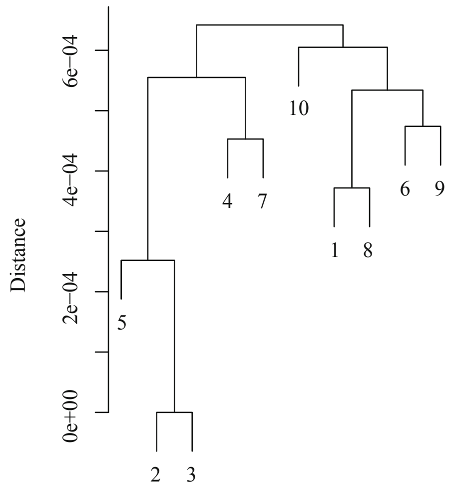


Fig. 3.16 Dendrogram corresponding to reduced modularity matrix in Fig. 3.15. (Adapted from Iino and Iyetomi 2012a)

It can be seen that there are strong relations between sub-communities within the same area. In the dendrogram, the sub-communities on the right side are mainly sub-communities located in the eastern part of Japan, such as Hokkaido and the Tohoku and north Kanto regions. On the other hand, the sub-communities on the left side are sub-communities located in the Chugoku, Shikoku, and Kyushu regions in western Japan. Tokyo bridges the two sides. Foods have expiration dates and some are produced locally, so firms in the food community may tend to establish business relations with their neighbors.

We carried out the same analysis for other communities and the results are shown in Table 3.3. They show that geographic distance and industrial closeness between sub-communities are important factors for understanding the strength of the interrelationship between sub-communities. Of course, for a more rigorous analysis of the interrelationship between sub-communities, more factors in addition to these two need to be taken into account.

3.7 Directional Bias in Interfirm Transactions

So far, our analysis has ignored the direction of the transaction relationships between firms, and it has been assumed that we are looking at an undirected network. In this section, we therefore examine the directional features of interfirm linkages at the level of the detected communities. The direction of each link in the transaction network is defined here in terms of the corresponding order flow. That is, if firm A (buyer) sends an order to firm B (supplier), the direction of the link between A and B is A $\rightarrow$ B.

The modularity formula (3.5) for an undirected network is readily generalized (Arenas et al. 2007; Leicht and Newman 2008) for the case of a directed network:

$$Q^{\mathrm{dir}} = \sum_{i=1}^{L} Q_i^{\mathrm{dir}} = \sum_{i=1}^{L} \left(e_{ii} - a_i^{\mathrm{in}} a_i^{\mathrm{out}}\right), \tag{3.14}$$

where e_{ii}, a_i^{in} and a_i^{out} are defined in terms of adjacency matrix A_{lm}, which takes value 1 if there is a link from firm l to firm m, and zero otherwise, as

$$e_{ii} = \frac{1}{M} \sum_{l \in V_i} \sum_{m \in V_i} A_{lm} \tag{3.15}$$

$$a_i^{\mathrm{in}} = \frac{1}{M} \sum_{l \in V_i} \sum_{m \in V_i} A_{lm} = \sum_{j} e_{ij} \tag{3.16}$$

$$a_i^{\mathrm{out}} = \frac{1}{M} \sum_{l \in V_i} \sum_{m \in V_i} A_{lm} = \sum_{j} e_{ji}. \tag{3.17}$$

We use the generalized modularity to extract communities by considering the direction of interfirm transaction relations. However, we find no appreciable differences between the two community structures obtained for the undirected and directed networks. This is demonstrated by the fact that Q^{dir} is 0.652 for the optimized community partition in the directed network and $Q^{\mathrm{dir}} = 0.655$ for that in the undirected network. The fact that these two values are very similar implies that link direction is not significant in determining community structure in the production network. The results in this section are thus based on the community structure obtained for the undirected network.

To quantify polarization of the link direction between communities, we introduce the polarization ratio:

$$P_{ij} = \frac{A_{ij} - A_{ji}}{A_{ij} + A_{ji}}, \tag{3.18}$$

where A_{ij} is the adjacency matrix between communities i and j, defined as the sum of A_{lm} over nodes l and m belonging to the two communities respectively:

$$A_{ij} = \sum_{l \in V_i} \sum_{m \in V_j} A_{lm}. \tag{3.19}$$

By definition, the polarization matrix is antisymmetric ($P_{ij} = -P_{ji}$). If the linkage between communities i and j is completely polarized, P_{ij} becomes ± 1 depending on its direction; if the linkage is evenly balanced, $P_{ij} = 0$.

We also attempt to visualize a transaction network with links that are directionally biased. Since the interaction force (3.1) does not reflect information on the link direction, we apply a uniform "magnetic field" $\boldsymbol{H}$ to the system to align directed links. The external field is presumed to exert an additional force on node l:

$$\boldsymbol{F}_{\mathrm{mag}} = \left(k_l^{\mathrm{in}} - k_l^{\mathrm{out}}\right) \boldsymbol{H}, \tag{3.20}$$

where k_l^{in} and k_l^{out} represent the in-degree (number of incoming links) and out-degree (number of outgoing links) of node l, respectively, imitating magnetic charges. The external force thus has the function of sorting nodes according to their positions in the supplier–buyer chain: buyers tend to be located on the upstream side and suppliers on the downstream side.

To examine directional bias between communities, we compute polarization (3.18) for all pairs of the top 10 major communities. Figure. 3.17 shows the results in matrix form. Further, Fig. 3.18 visualizes the coarse-grained network of communities with the accumulated "magnetic" force acting on community i:

$$\boldsymbol{F}_{\text{mag}}^{\text{c}} = \sum_{l \in V_i} \left(k_l^{\text{in}} - k_l^{\text{out}}\right) \boldsymbol{H}. \tag{3.21}$$

Each sphere in the figure represents a community, and the volume of each sphere is proportional to the number of nodes in the community. Arrow thickness between two spheres is proportional to the number of links between a pair of communities. In the visualization, the communities of the upper stream are configured in an upward direction and the lower-stream communities are configured in a downward direction by the uniform magnetic field $\boldsymbol{H}$. The third largest community, consisting mainly of food industry firms, occupies the uppermost position in the order stream among the major communities. This is because those firms tend to send orders to firms in the other communities. In particular, the tenth largest community, with many firms from the printing and paper products industry, receives a large share of its orders from the third largest community ($P_{3,10} = 0.675$). The second largest community, consisting of a group of construction firms in a metropolitan area, is also more upstream than most other communities, but it does have appreciable order flows from the third largest community. As shown in Fig. 3.17, the fourth and ninth largest communities, which mainly consist of local construction firms, have very polarized linkages with the second largest community ($P_{2,4} = 0.289$, $P_{2,9} = 0.237$). This shows that there tends to be a strong relationship between general contractors and subcontractors in the construction industry. Since the largest community has no substantially polarized linkages with the other communities except the third largest, it is located near the bottom.

3.8 Conclusion

This chapter examined the community structure of the Japanese production network, which covers virtually the entire economy of Japan. We started by visualizing the optimized configurations of the entire network and its partials in three-dimensional space. Doing so showed that the network has a very nonuniform structure. We then broke down the network into components at the sub-community level with the aid of a recursive community detection technique based on modularity, in which transaction relations were regarded as undirected links. A vast range of components were thus identified, and major communities and their internal structures were described in

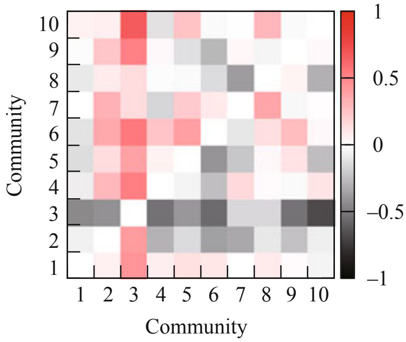


Fig. 3.17 Polarization matrix for the top 10 communities in the transaction network in Table 3.1. (Adapted from Iino and Iyetomi 2012b)

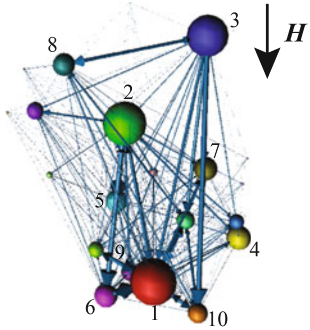


Fig. 3.18 Visualization of directed inter-linkages among the top 10 communities in the transaction network. (Adapted from Iino and Iyetomi 2012b)

terms of their regional, industry, and business affiliations. Furthermore, we measured the strengths of the relationships between (sub-)communities employing a reduced modularity matrix, and relations between the (sub-)communities were represented in the form of dendrograms. Finally, we examined the hierarchical relationship between the communities by computing the polarization ratio of orders flowing between them. Despite the fact that community detection was performed without using directional information of the orders, some of the relations between the communities were notably biased in direction.

Recently, economists (Acemoglu et al. 2012; Atalay et al. 2011; Cainelli et al. 2012; Luo et al. 2012) have also begun to recognize the importance of explicitly taking interfirm links into account when seeking to understand economic issues. These issues include the heterogeneity of firms, such as the fact that their size tends to follow a power-law distribution, the origin of business cycles, and the possibility of a chain-reaction in firm bankruptcies. We hope that the present study provides the impetus for further collaboration between physicists like us and economists, with fruitful results.

Acknowledgement This study received financial support through the *Program for Promoting Methodological Innovation in Humanities and Social Sciences by Cross-Disciplinary Fusing* from the Japan Society for the Promotion of Science (JSPS) and from the JSPS Grant-in-Aid for Scientific

Research No. 22300080. We thank the Research Institute of Economy, Trade and Industry (RIETI) for providing us with access to the data on transactions between Japanese firms. We are grateful to Hideaki Aoyama, Yoshi Fujiwara, Yuichi Ikeda, Wataru Souma, and Hiroshi Yoshikawa for useful discussions on the subjects herein.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica, 80*(5), 1977–2016.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74,* 47–97.

Allen, M. P., & Tildesley, D. J. (1987). *Computer simulation of liquids.* Oxford: Oxford University Press.

Arenas, A., Duch, J., Fernández, A., & Gómez, S. (2007). Size reduction of complex networks preserving modularity. *New Journal of Physics, 9,* 176.

Atalay, E., Hortaçsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *PNAS, 108,* 5199–5202.

Barabási, A.-L. (2002). *Linked: The new science of network.* Cambridge: Perseus Publishing.

Barnes, J., & Hut, P. (1986). A hierarchical O(N log N) force-calculation algorithm. *Nature, 324,* 446–449.

Battista, G. D., Eades, P., Tamassia, R., & Tollis, I. G. (1998). *Graph drawing: Algorithms for the visualization of graphs.* Englewood Cliffs: Prentice Hall.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistics Mechanics,* P10008.

Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski Z., & Wagner, D. (2008). On modularity clustering. *IEEE Transactions on Knowledge and Data Engineering, 20*(2), 172–188.

Buchanan, M. (2002). *NEXUS: Small worlds and the groundbreaking science of networks.* New York: W. W. Norton&Co.

Cainelli, G., Montresor, S., & Marzetti, G. V. (2012). Production and financial linkages in inter-firm networks: Structural variety, risk-sharing and resilience. *Journal of Evolutionary Economics, 22*(4), 711–734.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E, 70,* 066111.

Duch, J., & Arenas, A. (2005). Community detection in complex networks using extremal optimization. *Physical Review E, 72,* 027104.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports, 486,* 75–174.

Fortunato, S., & Barthélemy, M. (2007). Resolution limit in community detection. *PNAS, 104,* 36.

Frenkel, D., & Smit, B. (2002). *Understanding molecular simululation: From algorithms to applications* (2nd ed.). San Diego: Academic Press.

Fujiwara, Y., & Aoyama, H. (2010). Large-scale structure of a nation-wide production network. *The European Physical Journal B, 77,* 565–580.

Guimerà, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature, 433,* 895–900.

Hu, Y. (2006). Efficient, high-quality force-directed graph drawing. *The Mathematica Journal, 10,* 37–71.

Iino, T., & Iyetomi, H. (2011). Multiscale community analysis of a production network of firms in Japan. In J. Watada, et al. (Eds.), *Intelligent decision technologies, SIST 10* (pp. 537–545). Berlin: Springer.

Iino, T., & Iyetomi, H. (2012a). Subcommunities and their mutual relationships in a transaction network. *Progress of Theoretical Physics, 194,* 144–157.

Iino, T., & Iyetomi, H. (2012b). Directional bias between communities of a production network in Japan. In *Intelligent decision technologies, Vol. 2, SIST 16* (pp. 273–280). Berlin: Springer.
Iino, T., Kamehama, K., Iyetomi, H., Ikeda, Y., Ohnishi, T., Takayasu, H., & Takayasu, M. (2010). Community structure in a large-scale transaction network and visualization. *Journal of Physics: Conference Series, 221,* 012012.
Kamehama, K., Iino, T., Iyetomi, H., Ikeda, Y., Ohnishi, T., Takayasu, H., & Takayasu, M. (2010). Structure analyses of a large-scale transaction network through visualization based on molecular dynamics. *Journal of Physics: Conference Series, 221,* 012013.
Leicht, E. A., & Newman, M. E. J. (2008). Community structure in directed networks. *Physical Review Letters, 100,* 118703.
Luo, J., Baldwin, C. Y., Whitney, D. E., & Magee, C. L. (2012). The architecture of transaction networks: A comparative analysis of hierarchy in two sectors. *Industrial and Corprate Change, 21*(6), 1307–1335.
Medus, A., Acuña, G., & Dorso, C. O. (2005). Detection of community structures in networks via global optimization. *Physica A, 358,* 593–604.
Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45,* 167–256.
Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E, 69,* 066133.
Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 8577–8582.
Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69,* 026113.
Pfalzner, S. (1996). *Many-body tree methods in physics.* New York: Cambridge University Press.
Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E, 74,* 016110.
Watts, D. J. (1999). *Small worlds: The dynamics of networks between order and randomness.* Princeton: Princeton University Press.
Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393,* 440–442.

Takashi Iino is a Technical Staff at Faculty of Science, Niigata University (2007–). He is also working as a Research Assistant at the Research Institute of Economy, Trade and Industry (RIETI) (2013–). His fields of research are Network Science and Econophysics. Especially, he is doing research on community structures of real networks by taking advantage of state-of-the-art technology in information processing to deal with large-scale data. Also he is interested in visualization of complex networks using the molecular dynamics simulation method based on physical modeling. He received a B.A. and a Ph.D. both in physics from Niigata University.

Hiroshi Iyetomi is a Professor of Mathematics at Niigata University, working in econophysics, network science, and multivariate analysis. He first studied strongly coupled plasma physics as an Assistant Professor of Physics at the University of Tokyo (1984–1993). He then migrated to the current position via Researcher for semiconductor processing apparatuses at Energy Research Laboratory, Hitachi, Ltd. (1993–1994), Associate and later Full Professor for condensed matter and computational physics at Niigata University (1995–2012), and Project Professor for econophysics at the University of Tokyo (2012–2013). Also he studied condensed matter physics as a Research Associate at Argonne National Laboratory (1987– 1989), a Visiting Associate Professor at Louisiana State University (1996, 2001), and a Visiting Fellow at Delft University of Technology (2002). He received a B.A. and a Ph.D. both in physics from the University of Tokyo.

Chapter 4
Interfirm Networks in Manufacturing Industry Agglomerations in Japan: Evidence from Survey Data

Iichiro Uesugi

Abstract This chapter focuses on interfirm networks in Japan's major manufacturing agglomerations and conducts fact-finding analyses on the following three issues: (1) the nature of interfirm transaction relationships, including developments in such relationships over time; (2) firms' participation in network activities other than supplier-customer transactions; and (3) interactions between interfirm transaction relationships and relationships of other types. Based on the results of a unique firm-level survey completed by more than 1800 firms in December 2009, it is found, first, that the number of interfirm transaction relationships, especially those involving smaller firms, has declined over the past ten years. Second, the survey indicates that, apart from transaction relationships, many firms participated in group activities in individual industry associations and local chambers of commerce, indicating that firms tend to maintain relationships with firms similar to themselves. Third, it is found that bank lending attitudes are positively associated with the extent to which a firm is interconnected with other local firms, indicating that interfirm and firm-bank relationships are complementary.

Keywords Interfirm networks · Firm-bank relationships · Joint R&D

4.1 Introduction

Social and economic networks have been the subject of active research since the late 1990s. Comprehensive studies on recent developments in the economics of networks are provided, for example, by Jackson (2008), Goyal (2007), and Vega-Redondo (2007), highlighting, among other things, the following. Social and economic networks are composed of many types of agents, including households, individuals, firms, financial institutions, and a variety of relationships between these agents. Even if we limit our focus to firms and financial institutions, a wide variety of relationships can be observed. Examples include transaction relationships in which suppliers deliver goods and services to customers, research and development (R&D)

I. Uesugi (✉)
Institute of Economic Research, Hitotsubashi University, 2–1 Naka, Kunitachi, Tokyo, Japan
e-mail: iuesugi@ier.hit-u.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_4

collaborations, mutual equity holdings, and firms' borrowing relationships with banks.

Each of these types of relationships involves agents that have different economic motives. For example, suppliers and customers establish interfirm transaction relationships to diversify risk originating from demand-side uncertainty (Kranton and Minehart 2001), whereas firms often participate in joint R&D activities to share knowledge on cost reduction (Goyal and Moraga-Gonzalez 2001). Studies such as these provide us with a fair amount of theoretical knowledge on how the various types of interfirm networks are formed based on the underlying economic incentives of the participants. There is also a plethora of empirical studies that focus on interfirm and firm-bank relationships and examine their formation, as well as their impact on firm dynamics and economic efficiency. Several chapters in this book address these issues. For example, Ikeuchi and Okamuro (Chap. 8) examine the determinants of interfirm and firm-bank relationships, while Hosono and Miyakawa (Chap. 14) focus on the role of firm-bank relationships in transmitting shocks from banks to firms.

Nevertheless, our empirical knowledge on interfirm network formation is fragmented in several respects, mainly because of data limitations. First, some important pieces of information on interfirm transaction relationships are missing. These include information on the intensive margin of relationships, that is, information on the amount of goods and services transacted through these relationships and the evolution of such relationships over time. Second, there are many types of relationships involving firms other than existing transaction or collaborative R&D relationships which are not covered by existing databases. These include relationships in which new transaction partners join incumbent transaction networks or relationships consisting of group activities in organizations such as local chambers of commerce and individual manufacturing associations. Third, there is a relative lack of studies on interactions between different types of relationships or networks. One example is the interaction between firm-bank relationships and interfirm transaction relationships. In interfirm transaction networks, there may be hub companies whose default would damage supply chains and thereby cause negative externalities in the network. In this case, financial institutions that extend loans to many firms in such a network could also be negatively affected. To prevent any potential losses caused by such externalities, financial institutions may use their relationship with a hub firm threatened by default and rescue it, in which case there is an interdependence between interfirm transaction and firm-bank relationships.

Against this background, the present chapter focuses on interfirm networks in manufacturing agglomerations in Japan to examine the following three issues: (1) the nature of interfirm transaction relationships as well as developments in such relationships over time; (2) firms' participation in network activities other than supplier-customer transactions; and (3) interactions between interfirm transaction relationships and relationships of other types. To investigate these issues, the results of a unique firm-level survey implemented in December 2009 are used. A questionnaire was sent to more than 14,000 manufacturing firms located in Japan's three major industry agglomerations, namely the Keihin, Higashi-Osaka, and Hamamatsu areas. More than 1800 of these firms responded.[1]

The findings can be summarized as follows. Regarding the first issue, the survey results suggest that firms' number of transaction relationships tended to decrease rather than increase in the three major industry agglomerations over the past ten years, especially in the case of small firms and transaction relationships involving local customers and suppliers. This coincides with a substantial decline in the number of manufacturing firms located in these agglomerations and indicates that these industry agglomerations may have experienced negative feedback from the declining number of interfirm connections. Also, very few small manufacturing firms had direct transaction relationships with foreign firms. Given that previous studies such as Melitz (2003) suggest that whether firms begin exporting—and hence building transaction relationships with foreign firms—is closely related to their productivity, this finding suggests that the firms in the survey were either insufficiently productive, or there were other barriers such as low managerial capacity that prevented productive firms from exporting directly to foreign customers.

Regarding the second issue, the survey indicates that, apart from transaction relationships, many firms had established other types of interfirm links and participated in group activities. A large number of firms had joined activities in industry associations and local chambers of commerce. This suggests that regional proximity and shared industry interests are important determinants of firm links, which is reminiscent of the concept of "homophily" first proposed by Lazarsfeld and Merton (1954) referring to the fact that individuals tend to maintain relationships with those similar to themselves.

Regarding the third issue, finally, the survey suggests that bank lending attitudes appear to be affected by links among local firms. More specifically, hub customer firms for local suppliers are more likely to have loan applications approved than hub supplier firms for local customers. A possible interpretation of this result is that banks perceive the externalities associated with hub firms that purchase from local firms to be greater than those associated with hub firms that sell to local firms, so that banks are more accommodative in their lending attitudes toward the former. In addition, looking at the interaction between interfirm relationships of different types, the survey results suggest a positive association between interfirm transaction relationships and R&D collaboration. That being said, a substantial number of firms also engaged in R&D collaboration with firms other than transaction partners and with research institutions.

The remainder of this chapter is organized as follows. Section 4.2 provides an overview of the historical background of the three manufacturing agglomerations examined here as well as of the characteristics of interfirm transaction networks within each agglomeration in terms of a few key statistics calculated from a database

[1] A comprehensive summary of the survey, including the results regarding firm relocation choices and information exchange within interfirm networks, can be found in Uesugi et al. (2010; in Japanese).

collected and provided by Teikoku Databank (TDB), a Japanese credit research company. Section 4.3 outlines the survey implemented in December 2009 and presents an overview of the composition and basic characteristics of firms that responded. Next, Sections 4.4 through 4.6 detail the survey results and focus on various types of relationships in the industry agglomerations, namely, interfirm transaction relationships, other types of interfirm relationships and group activities, and firm-bank relationships, respectively. Section 4.7 then examines how different types of interfirm and firm-bank relationships interact. Finally, Section 4.8 provides some conclusions.

4.2 Japan's Three Major Manufacturing Agglomerations

The remaining part of this chapter focuses on three large manufacturing industry agglomerations in Japan. Specifically, three cities were picked: Higashi-Osaka in Osaka prefecture, Ota in Tokyo prefecture, and Hamamatsu in Shizuoka prefecture, which comprise the three largest agglomerations in terms of the number of manufacturing establishments in 2006.[2] For the analysis of Ota, we added the agglomerations of several adjacent cities, namely Shinagawa and Meguro in Tokyo prefecture, and Kawasaki, Yokohama, Sagamihara, and Yamato in Kanagawa prefecture. This was done to examine the extent of interfirm networks within a wider area. We label the area embracing these cities as Keihin. Thus, the following analysis focuses on the three industry agglomerations of Keihin, Higashi-Osaka, and Hamamatsu.[3] The geographic location of these agglomerations is shown in Fig. 4.1.

4.2.1 Historical Background

Each of the three industry agglomerations has its own unique history and characteristics, which are succinctly explained in SMEA (2010). The Keihin area, especially Ota city, is famous for the production of general machinery and fabricated metal products. Agglomeration started before World War II with a number of large manufacturing companies such as Mitsubishi Heavy Industries, Ebara Corporation, and Canon. In the 1960s, when a large number of factories were relocated because of spatial constraints, small subcontractors that had been supplying these factories stayed behind but changed their business strategy and started transacting with multiple customers rather than relying on a single customer. Through the formation of collaborative

[2] According to the 2006 *Establishment and Enterprise Census* published by the Japanese government, Higashi-Osaka, Ota, and Hamamatsu had 7388, 5953 and 5405 establishments, respectively, representing the top three among cities, special districts, and administrative districts in 2006.
[3] The following analysis uses firm-level survey data rather than establishment-level data. According to the 2006 *Establishment and Enterprise Census*, Keihin, Higashi-Osaka and Hamamatsu had 14,973, 2788, and 3220 manufacturing firms, respectively.

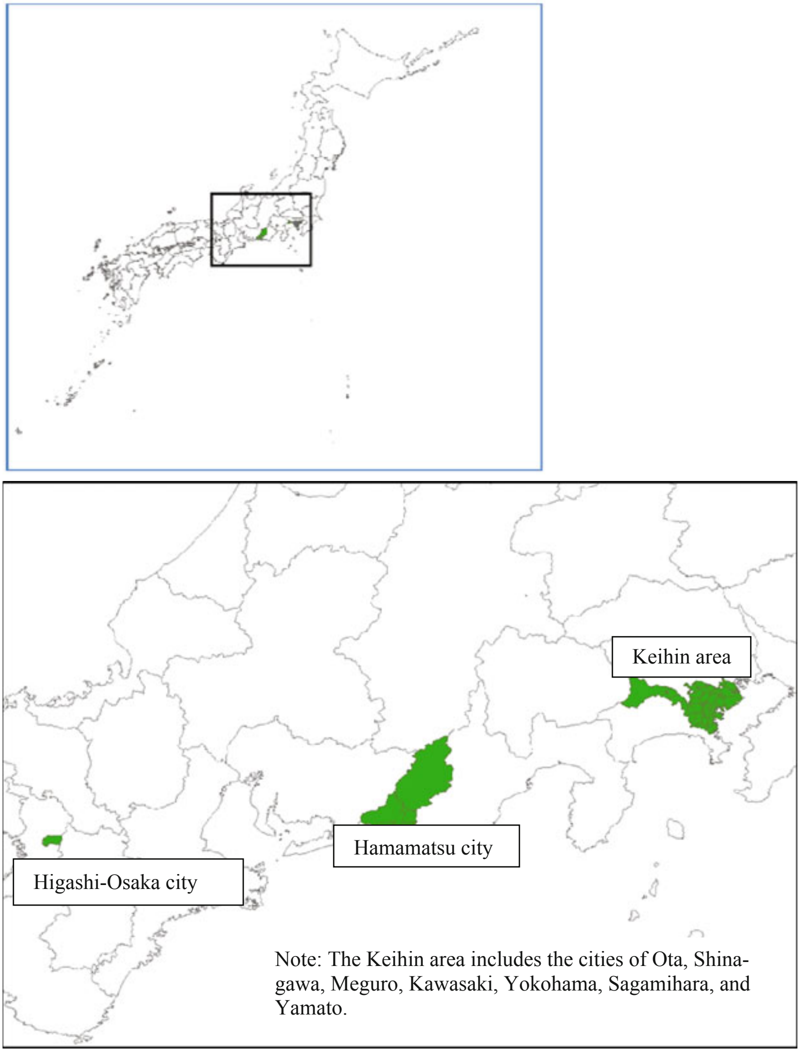


Fig. 4.1 Location of the three manufacturing agglomerations

networks comprising a large number of firms, the area developed a sophisticated production system able to supply a variety of high-precision parts and components at short notice. Adjacent cities, including Kawasaki and Yokohama in Kanagawa prefecture, became an increasingly important part of the industry agglomeration by receiving numerous manufacturing businesses relocated from the Tokyo area. Like Ota, these cities are home to a number of general machinery, fabricated metal products, and electronic appliances firms. The industry composition of the agglomeration has remained more or less unchanged over the past three decades.

Industry agglomeration in the Higashi-Osaka area started with textile, towel, metal net, and iron and copper wire production facilities. The agglomeration continued to grow as the area provided a convenient base for the relocation of manufacturing firms that used to be in the center of Osaka city, especially during Japan's high-speed growth era. Higashi-Osaka is heavily populated by firms involved in the fabricated metal products, general machinery, and plastic products industries. Today, the area is characterized by the production of a variety of final products with short delivery times.

Finally, the Hamamatsu area was originally famous for its cotton textile industry in the Edo era. Before World War II, there appeared several new manufacturing businesses, including musical instruments and machinery firms. After the war, a number of small manufacturers competed in the motorcycle industry, and several (Honda, Suzuki, and Yamaha) became leading manufacturing firms in the country. The first two, Honda and Suzuki, extended their businesses to automobile assembly. Overall, the dominant manufacturing industries in the area are the transportation machinery industry, followed by the general machinery and plastic products industries.

However, the three agglomerations have in common that the number of firms they comprise has been decreasing. Specifically, between 1996 and 2006, the number in the Keihin area dropped by 29 %, that in the Higashi-Osaka area by 28 %, and that in the Hamamatsu area by 27 %. These declines are in line with the decrease in the number of manufacturing firms in Japan overall, which was 29 % during the period.

4.2.2 *Interfirm Transaction Networks within the Agglomerations*

Before addressing the survey, let us take a look at the structure of local networks in the three agglomerations, using a dataset collected by TDB. TDB is one of the largest private credit information companies in the country and provides information on interfirm transaction relationships that identifies major transaction partners for a sizable number of manufacturing firms in the three agglomerations. TDB does not ask firms to report all their customers and suppliers, but only the major ones. Using information on these interfirm transaction relationships, it is possible to draw a transaction network graph for each agglomeration to get a sense of the interfirm relationships within them. There were 17,042 manufacturing firms registered in the three areas in the TDB database. Among these, information on firms' customers and suppliers was available for 11,047 firms. We focus on the largest network within each of the three industry agglomerations. Looking at the transaction links within each of the three industry agglomerations shows that the largest networks comprise more than half of all the firms in that agglomeration. Specifically, the largest network in the Keihin area encompassed 4774 firms, that in Higashi-Osaka 639 firms, and that in Hamamatsu 721.

Figure 4.2 and Table 4.1 examine these networks in more detail. In Fig. 4.2, each circle represents a firm, with the size of the circle indicating the size of the firm measured in terms of the number of employees. Note that the circles do not represent the

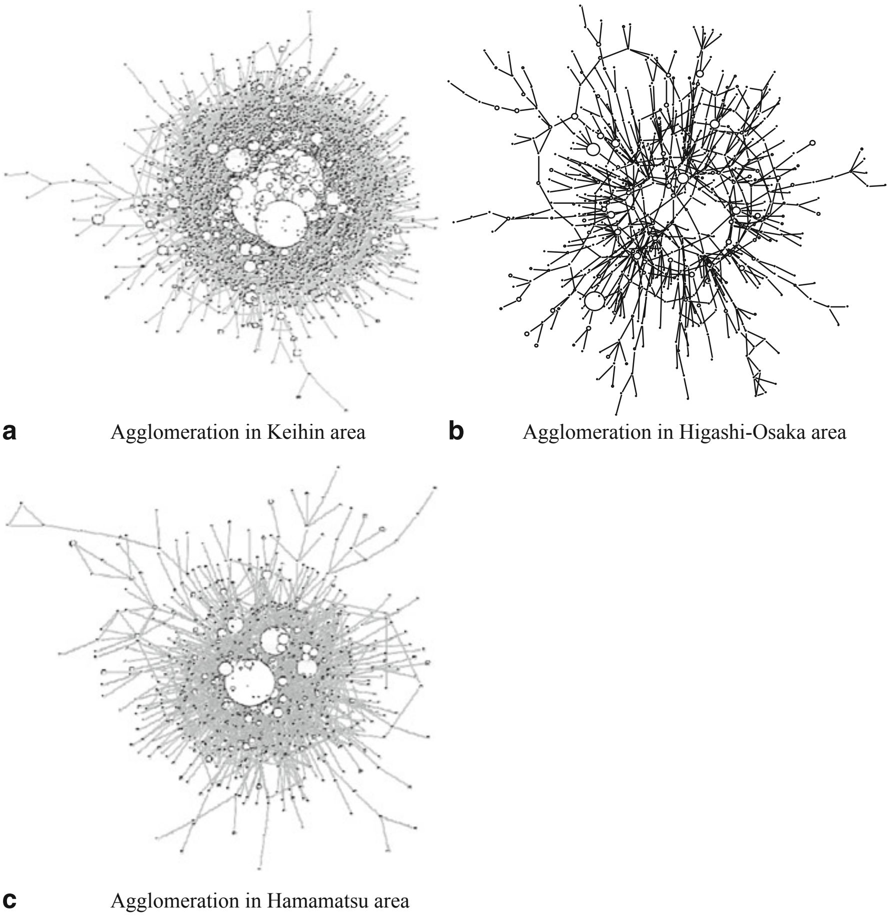


Fig. 4.2 Interfirm networks in the industry agglomerations **a** Agglomeration in Keihin area **b** Agglomeration in Higashi-Osaka area **c** Agglomeration in Hamamatsu area

actual geographical location of each firm. The figure indicates that in each of the agglomerations, firms are connected with each other through transaction relationships to form a large network. However, the way firms are connected differs notably among the three agglomerations. In the Keihin and Hamamatsu agglomerations, there are several large hub firms at the center of the network graphs. On the other hand, in Higashi-Osaka, large firms are not located in the center but instead are more at the periphery. At the same time, the density of links in the Keihin and Hamamatsu areas is much higher than in Higashi-Osaka, suggesting that the networks in the former two areas are much more tightly knit and, presumably, stronger.

These casual observations are supported by the network summary statistics calculated from the TDB database shown in Table 4.1, including the network statistics regarding measures of network centrality, such as degree, betweenness, closeness,

Table 4.1 Network statistics for the three agglomerations based on the TDB database

	Number of employees	Total number of degrees	Page rank	Betweenness (undirectional)	Closeness (undirectional)
Keihin					
Observations	4770	4774	4774	4774	4774
Mean	92.94486	4.136783	0.0002082	0.00088135	0.1980999
Std. Dev.	668.342	7.774149	0.0003054	0.00461746	0.0307226
Min.	0	1	0.0000604	0	0.0760735
Median	16	3	0.0001475	0.0001462	0.198286
Max.	30646	262	0.0103236	0.2081954	0.3134426
Higashi-Osaka					
Observations	639	639	639	639	639
Mean	43.77621	2.394366	0.0015649	0.0121846282	0.1177077
Std. Dev.	112.682	1.811068	0.0009567	0.0223554	0.0199126
Min.	0	1	0.0006276	0	0.0645161
Median	21	2	0.0013312	0.00313479	0.1181481
Max.	2231	15	0.0077086	0.21662716	0.1674541
Hamamatsu					
Observations	721	721	721	721	721
Mean	89.79889	4.601942	0.001387	0.00487528	0.2302695
Std. Dev.	586.6881	7.199996	0.0017052	0.01831672	0.0411702
Min.	0	1	0.0003949	0	0.0996816
Median	23	3	0.0009965	0.00071227	0.2342225
Max.	14266	104	0.0252456	0.37041303	0.374415

Degree is the number of transaction relationships for each firm, which is denoted by d_i.
Betweenness represents the importance of each node (firm) when connecting two different nodes (firms) in a network. It is defined as $\sum_{k \neq j : i \notin \{k,j\}} \frac{P_i(kj)/P(kj)}{(n-1)(n-2)/2}$
where $Pi(kj)$ is the number of shortest paths between k and j that i lies on, $P(kj)$ is the total number of shortest paths between k and j, and n is the number of firms in the network.
Closeness represents the average distance from a node (firm) to all the other nodes (firms) in a network. It is defined as $\frac{(n-1)}{\sum_{j \neq i} l(i,j)}$
where $l(i,j)$ is the number of links in the shortest path between i and j.
Page rank represents the importance of each node (firm) measured by the number of links with other nodes (firms) which have a high page rank themselves. It is calculated by solving the following equation for n firms in the network:
$Rank_i = (1-a) + a(Rank_{t1}/d_{t1} + \ldots + Rank_{tk}/d_{tk})$
where nodes (firms) $t1, \ldots, tk$ are those that have connections with node (firm) i and a is a damping factor that is customarily set at 0.85.
All the statistics above regard interfirm relationships as undirectional, that is, the direction of the flow of goods or services is not considered for the calculation of these statistics.

and page rank. These indicate that the industry network in Higashi-Osaka differs substantially from the networks in the other two agglomerations. First, in Higashi-Osaka, the average number of degrees is about 2.4, whereas in Keihin and Hamamatsu it is about 4.1 and 4.6, respectively. This indicates that firms in Higashi-Osaka are, on average, less connected with each other than those in the other two areas. Second, the distribution of the number of degrees shows that firms in Higashi-Osaka had a maximum of only 15 transaction relationships with other local firms, whereas firms in the other two areas had as many as 104 (Hamamatsu) and 262 (Keihin). Third, other measures of network centrality including page rank and closeness show that the distribution of the number of relationships was less skewed to the right for Higashi-Osaka firms than for those in the other two agglomerations. All this indicates that there are substantial differences across the agglomerations in terms of the way firms are connected. Specifically, in Higashi-Osaka, firms have a relatively similar number of links, resulting in the lack of a conspicuous hub structure; on the other hand, in Keihin and Hamamatsu, there are a limited number of firms that have considerably more links than other firms, meaning that they appear to act as hub firms.

That being said, the way TDB collects information on firms' transaction partners may result in an underestimation of the number of relationships of small firms in an agglomeration. This issue will be addressed again in Section 4.4.1 examining the distribution of all of firms' transaction relationships rather than only their major relationships.

4.3 Survey of Interfirm Relationships in the Three Major Industry Agglomerations in Japan

While the TDB data make it possible to gain some sense of the structure of networks, they provide very little other information on interfirm relationships. Therefore, in order to examine such relationships in the three major industry agglomerations in Japan, this study relies on a survey conducted specifically with this aim in mind. This section briefly introduces the structure of the survey and outlines the basic characteristics of responding firms.

4.3.1 Structure of the Survey

The survey of interfirm relationships in industry agglomerations was conducted in December 2009 by the Hitotsubashi Interfirm Network Research Project, together with the Small and Medium Enterprises Agency of the Ministry of Economy, Trade and Industry of Japan. The survey consisted of a questionnaire sent to the 14,094 manufacturing firms in the three industry agglomerations recorded in the TDB database. Given that, according to the 2006 *Establishment and Enterprise Census* carried out by the Japanese government, there were 20,981 manufacturing firms in the survey areas, the firms to which the questionnaire was sent make up a substantial share of the total population of firms in the areas. Out of the 14,094 firms that were sent the questionnaire, 1829 firms responded for a response rate of 13.0 %.

4.3.2 Basic Characteristics of Responding Firms

This subsection presents the basic characteristics of firms that responded to the survey, including their area, industry, age, number of employees, and performance. Table 4.2(a) shows the distribution of responding firms across the three areas and indicates that two-thirds were in the Keihin area, while about one-sixth each were located in the other two agglomeration areas.

Next, Table 4.2(b) shows the industry composition of firms in the sample. Almost all belong to the manufacturing sector, since the questionnaire was sent out only to firms identified as such in the TDB database. There are, however, some exceptions. Ninety-six responding firms were originally recorded as manufacturing firms in the database, but replied that they fell into the non-manufacturing sector. Among the manufacturing firms, 288 belonged to the fabricated metal products industry, 206 to other manufacturing, and 197 to the electronic appliances industry. Turning to Table 4.2(c), this shows the distribution of firm age as of December 2009. Firms aged between 31 and 50 years old made up the largest share, with those 31–40 years old accounting for 21 % and those 41–50 years old accounting for 19 %. Further, firm 10 years old or younger accounted for 8 %, while 6 % of firms were more than 70 years old.

Further, Table 4.2(d) shows the firm distribution in terms of the number of employees. Based on the legal definition of small and medium enterprises (SMEs) in Japan as firms with 300 or fewer employees, 1778 (99.3 %) of the firms were SMEs, while only 12 firms (0.7 %) were larger than that. This composition is roughly in line with the 2006 *Establishment and Enterprise Census*, where the corresponding values are 99.2 % and 0.8 %. Further, looking at the firm size distribution among SMEs, this shows that the share of very small firms with five employees or fewer was 29 % and that of small firms with 6–20 employees was 43 %. The corresponding values in the 2009 *Basic Survey on Small and Medium Enterprises* published by the Small and Medium Enterprise Agency are 48 and 31 %, respectively, meaning that the survey contained fewer very small firms and more small ones (6–20 employees or more).

Finally, Table 4.2(e) shows the development of firms' sales over the preceding year and the preceding three years. The questionnaire asked firms to choose from one of the following options to describe the development of their sales: decreasing, moderately decreasing, unchanged, moderately increasing, and increasing. Reflecting the fact that the survey was conducted in the midst of the severe recession following the collapse of Lehman Brothers, a large number of firms reported a decrease in sales for the year or the three years prior to December 2009, when the survey was conducted.

Table 4.2 Characteristics of respondent firms

(a) Location

	Observations	Share (%)
Keihin	1234	67.47
Higashi-Osaka	289	15.80
Hamamatsu	306	16.73
Total	1829	100

(b) Industry

	Observations	Share (%)
Fabricated metal products	288	16.04
Other manufacturing	206	11.48
Electronic appliances	197	10.97
Precision instruments and machinery	138	7.69
General machinery	130	7.24
Transportation equipment	108	6.02
Plastic products	108	6.02
Non-manufacturing	96	5.35
Printing and related	86	4.79
Food, beverage, and tobacco	67	3.73
Non-ferrous metal and its products	62	3.45
Textiles	60	3.34
Lumber and wood products	48	2.67
Electronic parts and devices	46	2.56
Iron and steel	36	2.01
Information and communication equipment	30	1.67
Chemical products	27	1.5
Pulp, paper, and paper products	24	1.34
Ceramic and pottery products	21	1.17
Rubber products	9	0.5
Petroleum and coal products	6	0.33
Leather products	2	0.11
Sum	1795	100

(c) Firm age

0–10 years	11–20	21–30	31–40	41–50	51–70	70+	Total
147	196	268	361	340	330	113	1755
8.38	11.17	15.27	20.57	19.37	18.8	6.44	100

Upper row shows the number of observations and lower row shows the share (%)

Table 4.2 (continued)

(d) Number of employees

0–5	6–20	21–50	51–100	101–300	301+	Total
514	762	329	119	54	12	1790
28.72	42.57	18.38	6.65	3.02	0.67	100

Upper row shows the number of observations and lower row shows the share (%)

(e) Development in the amount of sales during the one- and three-year periods preceding the survey

	Decreased	Decreased moderately	Stable	Increased moderately	Increased	Total
Preceding year	1081	313	204	126	65	1789
	60.42	17.5	11.4	7.04	3.63	100
Preceding three years	771	404	275	216	108	1774
	43.46	22.77	15.5	12.18	6.09	100

Upper row shows the number of observations and lower row shows the share (%)

4.4 Interfirm Transaction Relationships

The survey questionnaire asked firms to provide information on various types of relationships and links. This information includes interfirm transaction relationships, interfirm links other than commercial transactions such as R&D collaborations, and firm-bank relationships. The next three sections review each type of relationship or link.

To start with, the present section focuses on interfirm transaction relationships. Specifically, it examines the current situation regarding transaction relationships as well as trends over the one year and ten years preceding the survey. Since the survey was conducted in December 2009, the effects of the collapse of Lehman Brothers and the subsequent economic downturn are clearly visible over both time horizons in the trends in transaction relationships. Further, this section also looks at the nature of firms' transaction relationships with their major customers and suppliers. The survey questionnaire asked firms to report up to five major customers and suppliers. The analysis here focuses on firms' largest customer and supplier, since the characteristics of the second through fifth largest transaction partners appeared to be qualitatively similar to the characteristics of the largest partners.

4.4.1 Trends in Interfirm Transaction Relationships

Defining interfirm transaction relationships as relationships that last a certain period of time and possibly involve repeated transactions, let us start by looking at developments in these interfirm relationships over the past ten years. The questionnaire

Table 4.3 Evolution in the number of customers and suppliers over the past 10 years

	Decreased	Decreased moderately	Unchanged	Increased moderately	Increased	Total
Number of customers	300	323	606	350	174	1753
	17.11	18.43	34.57	19.97	9.93	100
Number of customers located in the same city	287	269	783	129	45	1513
	18.97	17.78	51.75	8.53	2.97	100
Number of suppliers	212	322	795	264	76	1669
	12.7	19.29	47.63	15.82	4.55	100
Number of suppliers located in the same city	208	299	828	124	29	1488
	13.98	20.09	55.65	8.33	1.95	100

Upper row shows the number of observations and lower row shows the share (%)

asked firms to report if the number of their customers and/or suppliers had decreased, moderately decreased, was unchanged, had moderately increased, or increased. It also asked the same with regard to the number of customers and suppliers located within firms' own city.[4] The reason that the questionnaire asked not for the exact number of customers and/or suppliers but rather whether the number had increased or decreased is that respondents probably would have found it difficult to remember exactly when and how many relationships over the past ten years had been established or terminated.

Table 4.3 shows the results. The share of firms reporting a decrease in the number of transaction partners was substantially larger than that of firms reporting an increase. Specifically, 36 % of firms reported a decrease in the number of customers over the previous ten years, while 30 % reported an increase. The difference was larger in the case of transaction relationships in the same city, which will be referred to as "local" transaction relationships, with 37 % reporting a decrease in the number of local customers and only 12 % reporting an increase. A similar pattern can be observed for supplier relationships. Specifically, 32 % of firms reported a decrease in the number of suppliers over the previous 10 years, while only 20 % reported an increase. Again, the difference is larger for local relationships, with 34 % of firms reporting a decrease in the number of local suppliers and 10 % reporting an increase.

Looking at the patterns for different firm sizes (not shown to conserve space) indicates that the overall pattern is largely shaped by smaller firms. That is, considerably more small firms saw a decrease in the number of transaction partners than an increase, while for larger firms, the shares are more or less balanced or even the

[4] Note that whereas the Higashi-Osaka and Hamamatsu agglomerations correspond to the cities of the same names, the Keihin agglomeration consists of the seven cities listed in Section 4.2. The survey questionnaire asked about customers/suppliers in the same city (rather than agglomeration), which may somewhat distort the results for Keihin using this question, given that firms in different but adjacent cities may be located very close to each other.

reverse. For instance, among firms with 0–5 employees, 50 % reported a decrease in the number of customers, while only 18 % reported an increase. In contrast, among firms with 101 + employees, 20 % reported a decrease in the numbers of customers, whereas 31 % reported an increase. A similar pattern was observed for the number of suppliers. Given that large firms inherently tend to have more transaction partners than small firms, the observed pattern suggests that the difference between large and small firms in terms of the number of transaction partners seems to have increased, with larger firms having more transaction partners and smaller firms having fewer transaction partners than ten years earlier.

Next, let us look at developments in the number of transaction relationships over the 1-year period preceding the survey. In this case, the questionnaire asked about the actual number of current transaction partners as well as the number of partners with which relationships were newly established or terminated. The results, presented in Table 4.4(a) and (b), indicate that a majority of firms established at least one new transaction relationship with a customer or supplier over the preceding year. Specifically, 78 % of firms reported that they established one or more new relationships with customers and 66 % with suppliers. The number of new customer relationships was larger than that of new supplier relationships. Similarly, a majority of firms reported that they terminated at least one transaction relationship with a customer and/or a supplier over the preceding year. Specifically, 66 % firms terminated at least one transaction relationship with a customer and 56 % with a supplier. Further, aggregating the number of relationships that were terminated by firms in the sample, more customer relationships were terminated than supplier relationships. Thus, summing up the patterns observed, in the case of both establishing and terminating transaction relationships, more firms reported doing so in the case of relationships with customers than with suppliers, suggesting that relationships with customers may be more fluid than those with suppliers.[5]

An interesting contrast between Tables 4.3 and 4.4(a) and (b) is that more firms reported a decrease in the number of transaction relationships than an increase over the preceding ten years, but over the preceding year, more established new relationships than terminated existing ones. Given that the year preceding the survey coincided with the deep recession following the collapse of Lehman Brothers in autumn 2008, a possible interpretation is that firms were trying to increase the number of transaction relationships as a short-run response; in the long run, however, firms lost customers and suppliers, possibly because of the declining number of firms not only in each industry agglomeration but also in the entire country.

[5] The aggregated number of new customer relationships and the number of new supplier relationships within a closed network should be the same if both suppliers and customers report the establishment of the same transaction relationships simultaneously. Therefore, the result that the number of new customer relationships is larger than that of new supplier relationships for firms in each of the agglomerations indicates that the transaction network is not closed in each of the agglomerations and that firms in these agglomerations are more likely to attract customers than they attract suppliers.

Table 4.4 Evolution in the number of customers and suppliers over the past 1 year

(a) Number of partners with which the firm newly started transactions

	0	1–2	3–6	7–10	11–30	31+	Total
All customers	247	317	295	120	81	66	1126
	21.94	28.15	26.2	10.66	7.19	5.86	100
Customers located in the same city	465	195	94	25	24	20	823
	56.5	23.69	11.42	3.04	2.92	2.43	100
All suppliers	340	308	196	74	46	38	1002
	33.93	30.74	19.56	7.39	4.59	3.79	100
Suppliers located in the same city	456	190	97	18	20	9	790
	57.72	24.05	12.28	2.28	2.53	1.14	100

Upper row shows the number of observations and lower row shows the share (%)

(b) Number of partners with which the firm terminated transactions

	0	1–2	3–6	7–10	11–30	31+	Total
All customers	368	378	216	57	34	15	1068
	34.46	35.39	20.22	5.34	3.18	1.4	100
Customers located in the same city	477	188	71	17	5	3	761
	62.68	24.7	9.33	2.23	0.66	0.39	100
All suppliers	436	368	131	33	15	1	984
	44.31	37.4	13.31	3.35	1.52	0.1	100
Suppliers located in the same city	469	222	46	9	2	0	748
	62.7	29.68	6.15	1.2	0.27	0	100

Upper row shows the number of observations and lower row shows the share (%)

(c) Number of current transaction partners

	0	1–3	4–10	11–30	31–100	101+	Total
All customers	5	146	355	410	352	227	1495
	0.33	9.77	23.75	27.42	23.55	15.18	100
Customers located in the same city	222	320	286	168	108	36	1140
	19.47	28.07	25.09	14.74	9.47	3.16	100
All suppliers	15	134	371	458	369	77	1424
	1.05	9.41	26.05	32.16	25.91	5.41	100
Suppliers located in the same city	147	304	371	234	85	10	1151
	12.77	26.41	32.23	20.33	7.38	0.87	100

Upper row shows the number of observations and lower row shows the share (%)

Table 4.4 (continued)

(d) Number of current transaction partners by location

All customers

	0	1–3	4–10	11–30	31–100	101+	Total
Keihin	3	92	253	278	235	133	994
	0.30	9.26	25.45	27.97	23.64	13.38	100
Higashi-Osaka	0	15	41	58	72	60	246
	0.00	6.10	16.67	23.58	29.27	24.39	100
Hamamatsu	2	39	61	74	45	34	255
	0.78	15.29	23.92	29.02	17.65	13.33	100

All suppliers

	0	1–3	4–10	11–30	31–100	101+	Total
Keihin	9	78	242	327	237	54	947
	0.95	8.24	25.55	34.53	25.03	5.70	100
Higashi-Osaka	1	16	45	77	73	17	229
	0.44	6.99	19.65	33.62	31.88	7.42	100
Hamamatsu	5	40	84	54	59	6	248
	2.02	16.13	33.87	21.77	23.79	2.42	100

Upper row shows the number of observations and lower row shows the share (%)

Lastly, Table 4.4(c) focuses on the number of current transaction relationships. Looking at the figures in detail, 15 % of firms had 101 + customer relationships, while only 5 % had 101 + supplier relationships. Comparing the number of firms' current customer and supplier relationships (not shown in the table), the median in both cases is 20, but the mean is 74 in the case of the former and 34 in the case of the latter; that is, firms, on average, tend to have more customer than supplier relationships. Further, comparing the distribution of customer relationships in the three agglomerations, Table 4.4(d) indicates that these are quite different. Firms in Higashi-Osaka had the largest numbers of customers and suppliers, with 24 % and 7 % having 101 + customers and suppliers, respectively. In contrast, firms in Hamamatsu and Keihin had smaller numbers of customers and suppliers, with 13 and 2 % having only 101 + customers and suppliers in Hamamatsu and 13 and 7 % in Keihin. Regarding the number of customers and suppliers located in the same city, a similar contrast is observed. That is, the distributions of the number of suppliers and customers are more skewed to the right in Higashi-Osaka than in Keihin (results not shown). That is, firms in Higashi-Osaka were more connected with each other than those in the other two agglomerations, especially Keihin.

The results of the survey seem to contradict those obtained from the TDB database described in Section 4.2.2, in which Higashi-Osaka had much sparser interfirm transaction relationships within its agglomeration than did Keihin and Hamamatsu. That is, in Table 4.1, the mean, median, and maximum number of transaction partners in

Higashi-Osaka is smaller than in Keihin and Hamamatsu, while the survey results discussed in the previous paragraph indicate that firms in Higashi-Osaka have more transaction partners than those in the other two agglomerations. However, this apparent contradiction can be attributed to differences in the way the survey and the TDB database count the number of transaction partners. The survey asked firms to report the number of all suppliers and customers whereas the TDB database collects information only on firms' major suppliers and customers. The latter database therefore understates the number of relationships involving small firms, because in many cases they are not one of the major transaction partners of the reporting firm. This problem is likely to be more severe in agglomerations composed mostly of small firms since the TDB database fails to detect connections by these firms. In sum, comparing the three agglomerations, the evidence in Section 4.2.2 above and this section indicates that firms in Higashi-Osaka tended to have the largest number of interfirm transaction relationships, mainly because Higashi-Osaka is characterized by a large number of links among small firms as opposed to a small number of links involving large firms.

4.4.2 Firms' Relationship with Primary Transaction Partners

The next point of interest is firms' relationship with their major customers and suppliers. The questionnaire asked firms to provide details on the characteristics of their top five customers and suppliers (such as location and size), as well as their relationship with them, including the length of the relationship and the share of the bilateral transaction amount with these major customers or suppliers in the total amount of transactions with all the customers or suppliers. The analysis here focuses only on firms' largest customer and/or supplier, which will be referred to as firms' primary customer and supplier. Tables 4.5(a) through (c) respectively show the survey results for primary transaction partners' location and size as well as the length of transaction relationships with them.

Starting with the location of the primary customer's headquarters, Table 4.5(a) shows that in many cases these were in the same prefecture as those of the responding firms. Looking at the different industry agglomerations, the share of firms whose primary customer was located in the same prefecture is 40 % in Keihin, 50 % in Higashi-Osaka, and 64 % in Hamamatsu.

The table also shows that very few firms in the three agglomerations had direct customer relationships with firms with headquarters abroad. In fact, in Keihin there were only nine such firms, while in Higashi-Osaka and Hamamatsu there were none at all. It is surprising that hardly any manufacturing firms in technologically-advanced industry agglomerations had established direct customer relationships with firms abroad.[6]

[6] That being said, it is possible that responding firms had established direct customer relationships with the affiliates of foreign firms in Japan, or with Japanese firms that had direct foreign customer relationships, but the survey did not explicitly ask about this.

Table 4.5 Characteristics of primary customers and suppliers

(a) Location

Primary customer	Keihin	Higashi-Osaka	Hamamatsu
Same prefecture	331	80	127
	40.12	49.69	64.47
Different prefecture	494	81	70
	59.88	50.31	35.53
Abroad	9	0	0
	1.09	0.00	0.00
Total	825	161	197
	100	100	100
Primary supplier	Keihin	Higashi-Osaka	Hamamatsu
Same prefecture	352	90	130
	46.62	62.07	69.15
Different prefecture	380	49	57
	50.33	33.79	30.32
Abroad	23	6	1
	3.05	4.14	0.53
Total	755	145	188
	100	100	100

Upper row shows the number of observations and lower row shows the share(%)

(b) Number of employees

Primary customer							
1–5	6–20	21–50	51–100	101–300	301–1000	1001+	Total
64	122	155	128	227	210	362	1268
5.05	9.62	12.22	10.09	17.9	16.56	28.55	100
Primary supplier							
1–5	6–20	21–50	51–100	101–300	301–1000	1001+	Total
187	306	201	148	135	79	84	1140
16.4	26.84	17.63	12.98	11.84	6.93	7.37	100

Upper row shows the number of observations and lower row shows the share (%)

Table 4.5 (continued)

(c) Length of relationship

Primary supplier							
Less than 2 yrs	2–5 yrs	5–10 yrs	10–15 yrs	15–20 yrs	20–30 yrs	30+ yrs	Total
47	75	190	146	195	248	402	1303
3.61	5.76	14.58	11.2	14.97	19.03	30.85	100
Primary supplier							
Less than 2 yrs	2–5 yrs	5–10 yrs	10–15 yrs	15–20 yrs	20–30 yrs	30+ yrs	Total
30	97	199	162	199	231	271	1189
2.52	8.16	16.74	13.62	16.74	19.43	22.79	100

Upper row shows the number of observations and lower row shows the share (%)

The patterns are quite similar in terms of the geographic locations of primary suppliers; that is, in many cases, firms' primary supplier was located in the same prefecture as the responding firm. Specifically, the share of firms for which this was the case was 47 % in the Keihin, 62 % in the Higashi-Osaka, and 69 % in the Hamamatsu agglomeration. And again, very few firms in these agglomerations had a direct relationship with suppliers located abroad. The number of firms for which this was the case was 23 in Keihin, 6 in Higashi-Osaka, and 1 in Hamamatsu. These numbers were larger than in the case of primary customers. Although these numbers are larger than in the case of primary customers, they are still extremely small, especially in the case of the Hamamatsu area.

Next, Table 4.5(b) examines the size of firms' primary customer, measured in terms of their number of employees. This shows that customer firms tend to be very large, given that nearly half of primary customers had more than 300 employees. Comparing the figures in Table 4.5 with those for the size of respondents themselves in Table 4.2(d) indicates that the size distribution of primary customers was much more skewed to the right, given that only 1 % of respondents had more than 300 employees. Similarly, respondent firms' primary supplier tended to be larger than the respondents themselves, although, with 14 %, the share of suppliers with more than 300 employees was considerably smaller than the share of customers of that size.

Finally, Table 4.5(c) shows that the length of the relationship with the primary customer typically was very long: more than 30 % of responding firms reported relationships of more than 30 years. There was a similar but weaker pattern for the length of the relationship with the primary supplier: 23 % of firms reported a relationship lasting more than 30 years. There is an interesting contrast between Table 4.5(c) here and Tables 4.4(a) and (b) above. That is, although firms established or terminated a considerable number of customer and supplier relationships in the one-year span preceding the survey, a substantial share of them had maintained a relationship with their primary supplier for more than 30 years. It therefore appears that firms clearly distinguish between core transaction partners with whom they maintain a long-term relationship and peripheral ones with whom they frequently start and terminate relationships.

Table 4.6 Number of firms participating in network activities other than interfirm transactions (multiple choice)

Local chamber of commerce	Industry association	Inter-industry study groups	Allocation of orders	Introduction of customers/ suppliers	Customer-supplier matching meetings	Joint R&D	Total
551	606	515	92	305	266	268	1156
47.66	52.42	44.55	7.96	26.38	23.01	23.18	100

Upper row shows the number of observations and lower row shows the share (%)

4.5 Other Types of Interfirm Relationships and Group Activities

This section considers the survey results on other types of activities through interfirm relationships and groups. The activities include not only bilateral ones, such as joint R&D or the introduction of new customers and suppliers but also ones in which many firms participate, such as activities in local chambers of commerce and manufacturing industry associations. The survey questionnaire asked firms about the types of activities other than transactions they engaged in in bilateral relationships and about group activities that they engaged in.

Table 4.6 presents the results, showing the share of firms that had joined the local chamber of commerce, an industry association, or an inter-industry study group, that allocated orders among firms, that used transaction relationships for the introduction of new customers or suppliers, that participated in meetings for customer-supplier matching, and that engaged in joint R&D activities. The results indicate that 52 % of responding firms participated in an industry association, 48 % in a local chamber of commerce, and 45 % in inter-industry study groups. Further, 26 % introduced new customers/suppliers through transaction networks, 23 % participated in meetings for supplier-customer matching, 23 % took part in joint R&D, and 8 % participated in the allocation of orders. However, although not shown in the table, the share of firms that participated in a particular activity varied greatly depending on firm size and age. Participation rates in local chambers of commerce and industry associations were substantially higher among larger firms, while participation in joint R&D projects was greater among younger firms. On the other hand, participation rates in other activities did not depend on firm size or age.

4.6 Firm-Bank Relationships

Next, this section looks at the relationship between firms and financial institutions with which responding firms transacted. The survey questionnaire asked firms about their relationship with their two largest lenders. Specifically, it asked about the lender type, the share of each lender in the firm's total outstanding loans, the length of the relationship with the lender, and the lending attitude. The analysis here focuses on the relationship with the primary financial institution, that is, the financial institution from which the firm had borrowed the most at the time of the survey.

The results are presented in Table 4.7, with panel (a) showing what type of financial institution the primary lender is. In Japan, there are various types of financial institutions, which can be categorized in terms of their size and legal status. Starting with the major, regional, and second-tier regional banks, these are all regulated by the same legal codes but differ in size. Major banks are the largest, followed by regional and second-tier regional banks.[7] Next, *shinkin* banks and credit cooperatives are in most cases smaller than banks in terms of their deposits and are regulated by legal codes that differ from those for the first three types of banks. Finally, there are several government-affiliated banks, including the Japan Finance Corporation and Shoko Chukin Bank. In the sample, 42 % of responding firms indicated that their primary bank was a *shinkin*, 25 % indicted that it was a major bank, 19 % that it was a regional bank, 10 % that it was a government financial institution, and 2 % each that it was a credit cooperative or a second-tier regional bank.

Table 4.7(b) shows the share that loans extended by the primary financial institution make up in a firm's total amount of outstanding loans. The results indicate that for 31 % of firms, the primary bank accounted for more than 80 % of their borrowing, and for 77 % of firms, it accounted for at least 40 %. These figures indicate that even though firms in the survey have relationships with one or more other banks, they tend to rely heavily on their primary financial institution.

Table 4.7(c) shows the distribution of the length of firms' relationship with their primary bank, which indicates that such relationships tend to be long-lasting. More than half (54 %) of the firms reported that their relationship with their primary bank exceeded 20 years, while only 10 % reported a relationship of no more than five years. The average and median relationship lengths were slightly greater than in the case of firms' relationship with their largest customer and/or supplier.

Finally, Table 4.7(d) focuses on the primary bank's lending attitude. Firm were given the following choices to describe the bank's lending attitude when the firm filed a loan application in the preceding year: application rejected, approved but for a lower amount than requested, fully approved, loans offered in addition to the original application, and no applications. The results indicate that 60 % of firms received approval for the full loan amount they requested or were even offered additional loans.

4.7 Interactions Between Relationships of Different Types

This section examines the interactions between different types of interfirm or firm-bank relationships. It starts by focusing on firm-bank relationships to see if any interactions with interfirm transaction relationships can be found, and

---

[7] There are several trust banks among the major banks, whose fiduciary businesses are regulated by different legal codes from the Banking Act. These banks are regarded as distinct from other major banks. However, only very few firms in the sample had a trust bank as their primary financial institution.

Table 4.7 Relationship with the primary bank

(a) Firms' primary bank by type

Major bank	Regional bank	Second-tier regional bank	Shinkin bank	Credit cooperative	Government financial institution	Total
347	264	23	573	27	136	1370
25.33	19.27	1.68	41.82	1.97	9.93	100

Upper row shows the number of observations and lower row shows the share (%)

(b) Share of the primary bank in terms of loan amount outstanding

−20%	20−40%	40−60%	60−80%	80%+	Total
103	144	271	216	327	1061
9.71	13.57	25.54	20.36	30.82	100

Upper row shows the number of observations and lower row shows the share (%)

(c) Length of relationship with the primary bank

Less than 2 yrs	2−5 yrs	5−10 yrs	10−15 yrs	15−20 yrs	20−30 yrs	30+ yrs	Total
34	100	193	123	175	250	473	1348
2.52	7.42	14.32	9.12	12.98	18.55	35.09	100

Upper row shows the number of observations and lower row shows the share (%)

(d) Lending attitude of the primary bank

Rejected	Approved but lower amount than requested	Approved full amount	Offered additional loans	Did not apply for loan	Total
55	91	627	166	391	1330
4.14	6.84	47.14	12.48	29.4	100

Upper row shows the number of observations and lower row shows the share (%)

then investigates how interfirm research collaborations are related to transaction relationships.

4.7.1 Firm-Bank Relationships and Interfirm Transaction Relationships

Interfirm relationships may affect firm-bank relationships when the firm in question is a hub firm whose activities affect other firms in the network and banks have a strong interest in maintaining a lending relationship with such a firm. To examine this point, Table 4.8 compares bank lending attitudes toward hub and peripheral firms, where hub supplier firms are firms that have a large number of relationships with customer firms and hub customer firms are firms that have a large number of relationships with

Table 4.8 Primary bank lending attitude toward hub and peripheral firms

Number of customers in the same city	Rejected	Approved but lower amount than requested	Approved full amount	Offered additional loans	Did not apply for loan	Total
Top quartile=Hub supplier firms	9	14	124	29	53	229
	3.93	6.11	54.15	12.66	23.14	
Bottom quartile=Peripheral supplier firms	17	18	128	30	75	268
	6.34	6.72	47.76	11.19	27.99	
Number of suppliers in the same city	**Rejected**	**Approved but lower amount than requested**	**Approved full amount**	**Offered additional loans**	**Did not apply for loan**	**Total**
Top quartile=Hub customer firms	7	15	127	33	52	234
	3.06	6.55	55.46	14.41	22.71	
Bottom quartile=Peripheral customer firms	14	18	117	23	88	260
	5.22	6.72	43.66	8.58	32.84	

Upper row shows the number of observations and lower row shows the share (%)

supplier firms. Specifically, hub supplier/customer firms are defined as firms in the top quartile in terms of the number of customer/supplier relationships in the same city, and peripheral ones as those in the bottom quartile.

Looking at the results, the table shows that hub supplier/customer firms were more likely to have the full amount of their loan request approved and to be offered additional loans than peripheral firms. Further, comparing the results for hub supplier firms and hub customer firms, the favorable treatment of hub firms vis-à-vis peripheral firms was more pronounced in the case of hub customers. Specifically, 67 % of hub supplier firms had the full loan amount approved or were offered additional loans compared to 59 % of peripheral supplier firms, for a difference of 8 percentage points. In contrast, in the case of hub and peripheral customer firms, the figures were 70 % and 52 % respectively, for a difference of 18 percentage points. These figures indicate that hub customer firms that purchase goods and services from local suppliers were treated better by their banks than hub supplier firms for local purchasers. This is consistent with the conjecture that hub customer firms are indispensable for local suppliers, so that if they were to default this would potentially result in substantial negative externalities, leading banks to maintain stronger relationships with these firms and be more forthcoming in providing funding. In contrast, hub firms for local purchasers may be more easily substituted and thus the negative externalities that would be brought about by their default would be relatively small.

4.7.2 Interaction between R&D Collaborations and Transaction Relationships

Interfirm R&D collaboration on the one hand and interfirm transactions in goods and service on the other are very different in nature, but they can complement each other. For instance, firms purchasing intermediate products often ask their suppliers to improve product quality or develop new products, providing the impetus for the customer and supplier to start collaborating on R&D. If such R&D collaboration is important, one would expect collaborative R&D relationships and interfirm transaction relationships to be positively related. That is, one would expect a substantial share of firms that have establish collaborative R&D relationships to also report transaction relationships with those R&D collaborators. On the other hand, studies suggest that other aspects apart from existing transaction relationships are important for collaboration. Branstetter and Sakakibara (2002), for instance, showed that technological proximity is important for the success of R&D collaborations, while Miotti and Sachwald (2003) argue that the choice of collaborator depends on the type of complementary resource required for innovation, in which case firms establish joint R&D activities not only with transaction partners, but also with other firms.

One of the survey questions examined this issue by asking firms about the characteristics of their R&D collaborators. Table 4.9(a) summarizes the results, showing that 34 % of responding firms indicated that their customers were R&D collaborators and 27 % that their suppliers were R&D collaborators. In contrast, only 20 % reported that the collaborators were firms other than customers or suppliers, indicating that interfirm transaction relationships are a dominant factor in determining the establishment of R&D collaborations. However, this does not necessarily mean that firms carry out joint R&D with only one type of collaborator. Table 4.9(b) shows the distribution of firms in terms of the number of different types of research collaborators they work with, including suppliers, customers, firms belonging to the same industry, other firms, universities, and research institutions. About a quarter of responding firms reported only one type of collaborator, but a larger number reported more than one type. This suggests that there is a certain amount of heterogeneity among partners involved in joint R&D activities.

4.8 Conclusions

This chapter addressed interfirm and firm-bank relationships in Japan's three major manufacturing agglomerations, examining developments in the number of transaction relationships over time, firms' participation in networking and industry activities other than interfirm transactions, and the interactions between interfirm transaction relationships and relationships of other types. The results of the survey conducted to investigate these issues can be summarized as follows.

Regarding developments in the total number of interfirm transaction relationships, more firms reported a decrease than an increase over the ten years preceding the

Table 4.9 R&D collaborators by type

(a) Type of R&D collaborator(s) (multiple choice)

Supplier(s)	Customer(s)	Firm(s) in the same industry	Other firm(s)	University	Public research institution(s)	Total
486	614	229	137	139	114	1829
26.57	33.57	12.52	7.49	7.60	6.23	100

Upper row shows the number of observations and lower row shows the share (%)

(b) Number of types of a firm's R&D collaborators

0	1	2	3	4	5	6	Total
922	433	300	123	39	14	4	1829
50.41	23.67	16.40	6.72	2.13	0.77	0.22	100

Upper row shows the number of observations and lower row shows the share (%)

survey. This was especially the case for small firms and transaction relationships involving local customers and suppliers. It has long been argued that low startup ratios and the resulting decline in the number of firms pose a serious problem to the economy. The results show that the substantial decline in the number of firms goes hand-in-hand with a decline in the number of interfirm links and therefore provide evidence of how the decreasing population of firms adversely affects the Japanese economy. However, although the number of interfirm relationships has been falling over the long-run, it increased in the year preceding the survey, probably in response to the economic downturn following the failure of Lehman Brothers. The results further indicated that very few firms in the three agglomerations had direct transaction relationships with foreign firms.

Next, looking at links other than bilateral transaction relationships, it was found that a large number of firms had joined industry associations and local chambers of commerce. Finally, it was examined whether firms' position within an agglomeration—i.e., whether they were hub or peripheral firms—affected bank lending attitudes. It was found that firms that were hub customers for local suppliers were more likely to have their loan applications approved than firms that were hub suppliers. A possible interpretation of this finding is that the externalities associated with hub firms purchasing from local firms are greater than those associated with hub firms selling to local firms, and that bank lending attitudes are accommodative in order to maintain lending relationships with these hub customer firms. Finally, it was found that a large number of firms that conducted joint R&D also had a transactional relationship with their collaborator.

An important caveat regarding all these findings is that they are based on simple summary statistics, not on multivariate regression analyses. The next step therefore is to conduct such analyses while keeping, for example, the possible endogeneity of variables in mind. Such analyses should help to provide a more accurate sense of the determinants of interfirm relationships and network structures and to discover the interrelationship among the different variables discussed here as well as among different types of interfirm links.

References

Branstetter, L. G., & Sakakibara, M. (2002). When do research consortia work well and why? Evidence from Japanese panel data. *American Economic Review, 92*(1), 143–159.

Goyal, S. (2007). *Connections*. Princeton: Princeton University Press.

Goyal, S., & Moraga-Gonzalez, J. L. (2001). R & D networks. *RAND Journal of Economics, 32*(4), 686–707.

Jackson, M. O. (2008). Social and economic networks. Princeton: Princeton University Press.

Kranton, R. E., & Minehart, D. F. (2001) A theory of buyer-seller networks. *American Economic Review, 91*(3), 485–508.

Lazarsfeld, P. F., & Merton, R. K. (1954) Friendship as a social process: A substantive and methodological analysis. In M. Berger (Ed.), *Freedom and control in modern society*. New York: Van Nostrand.

Melitz, M. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica, 71,* 1695–1725.

Miotti, L., & Sachwald, F. (2003). Co-operative R & D: Why and with whom? An integrated framework of analysis. *Research Policy, 32,* 1481–1499.

Small and Medium Enterprises Agency of Japan. (2010). *White paper on SMEs*. Tokyo: Gyosei

Uesugi, I. H., Okamuro, H. H., & Ohkubo, S. (2010). Interfirm networks in manufacturing agglomerations: A summary of the survey results implemented in December 2009 (Seizougyo shuhsekichi ni okeru kigyokan no tsunagari). Design of Interfirm Network to Achieve Sustainable Economic Growth WP Series No. 2 (in Japanese).

Vega-Redondo, F. (2007). *Complex social networks* (*Econometric society monographs*). Cambridge: Cambridge University Press.

Iichiro Uesugi is a professor at the Institute of Economic Research of Hitotsubashi University, Japan. He received his M.A. and Ph.D. in economics in 2000 from the University of California, San Diego. Prior to becoming a member of the faculty of Hitotsubashi University in 2011, he joined the Ministry of Economy, Trade and Industry (METI) in 1993 and worked in the Economic Policy Unit of the Minister's Secretariat, Macroeconomic Affairs Division; the Planning Division of the Coal Department in the Agency of Natural Resources and Energy; the Research Division of the Small and Medium Enterprises Agency; and the Research Institute of Economy, Trade and Industry. His research interests focus on firm dynamics, financial intermediation, and the Japanese economy. His research has been published in a number of scholarly journals including the *Journal of Money, Credit, and Banking*, the *Journal of Banking & Finance*, and the *Journal of the Japanese and International Economies*.

Part II
Networks, Economic Geography, and Firm Activities

Chapter 5
Economic Geography and Interfirm Transaction Networks

Kentaro Nakajima

Abstract Recent years have seen a remarkable expansion of empirical research using microdata on interfirm transactions that capture disaggregated firm-level transaction relationships. This chapter reviews the background to and recent empirical research on this issue. The chapter starts with a review of studies examining transaction networks without the use of transactions microdata and then shows the limitations of such research. Following this, a definition of transactions microdata and specific examples are provided. Finally, recent empirical research using transactions microdata is reviewed. The review focuses on the following three important questions. How substantial are geographic frictions in the formation of interfirm transaction networks? How great is the impact of networks on the behavior of agents within such networks? And to what extent does a geographically localized shock propagate through such networks? Finally, the future outlook in this line of research is considered.

Keywords Agglomeration · Supply chain networks · Gravity equation · Shock propagation

5.1 Introduction

Economic activities are not evenly distributed and there is a strong tendency for them to be concentrated in very small areas. From a global perspective, economic activities are concentrated in developed countries, but even within developed countries, they are not evenly distributed. For example, economic activities in Japan are concentrated in large metropolitan areas such as Greater Tokyo, which is home to more than one-quarter of the country's population. In fact, not only do economic activities overall tend to be concentrated, activities in particular industries also are often concentrated in so-called industrial clusters. For example, in the United States, the information

K. Nakajima (✉)
Faculty of Economics, Tohoku University, 27-1 Kawauchi Aoba-ku,
Sendai, Japan
Tel.: +81-22-795-6305
e-mail: knakajima@econ.tohoku.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_5

technology (IT) industry is concentrated in Silicon Valley and the auto industry in Detroit. In Japan, small and medium-sized manufacturers cluster in Ota Ward in Tokyo and Higashi-Osaka in Osaka Prefecture, while Sabae in Fukui Prefecture is the production center for the bulk of domestic optical glass. Given this backdrop, there has been a wave of research in recent years aimed at quantitatively identifying industry agglomeration in the economy. For example, Duranton and Overman (2005) showed that, as of 2000, around half of industries in the United Kingdom were agglomerated. Similarly, Nakajima et al. (2012a) found that, as of 2005, around half of the industries in Japan were agglomerated.

Theoretical analysis of the causes of agglomeration can be traced to Marshall (1890), who argued that the agglomeration of firms and workers gives rise to positive externalities. Examples of such positive externalities are the spillover of specialized knowledge, technology transfer, and the pooling of skilled workers within clusters. Another important aspect highlighted by Marshall is the role of input sharing. The clustering of firms supports auxiliary industries that produce intermediate goods. The proximity of firms that share certain common production processes increases their profits by reducing transaction costs such as costs associated with transportation and communication. Studies on urban agglomeration from this perspective include Borukhov and Hochman (1977), O'Hara (1977), and Fujita and Ogawa (1982). These studies explicitly introduce communication with other firms as an input for production and show that communication between firms is a driving force underlying the concentration of firms. Further, new economic geography (NEG) (see, e.g., Fujita et al. 1999) suggests that the transportation costs of goods determine the spatial distribution of economic activities. In a class of NEG models (see, e.g., Krugman and Venables 1995) that examine intermediate goods transactions, transportation costs in interfirm transactions play a key role in determining the geographic distribution of economic activities.

Thus, issues related to interfirm transactions are extremely important in determining the spatial distribution of industries. In addition to theoretical studies such as the ones mentioned above, there are numerous empirical studies examining interfirm transactions.

In recent years, microdata on interfirm transactions providing information on who transacts with whom have become increasingly available. Combined with the development of network theory (see, e.g., Goyal 2007; Jackson 2008), this has led to a spate of empirical studies using microdata that have provided new insights into interfirm transactions and the geographic distribution of transaction relationships and economic activities. The aim of this chapter is to provide a survey of such studies on the role of interfirm transactions in the spatial distribution of economic activities using microdata on interfirm transactions.

The remainder of this chapter is organized as follows. The next section looks at studies using aggregate data and addresses their limitations, while Section 6.3 provides examples of microdata on interfirm transactions. The following sections then present a review of studies using transactions microdata in three research areas. Specifically, Section 6.4 provides an overview of studies on the role of geographic frictions in giving rise to the formation of interfirm transaction networks. Section 6.5 then focuses on studies examining how networks affect the behavior of members within such networks, while Section 6.6 addresses research on the propagation of a geographically confined shock through networks. Section 6.7 concludes.

5.2 Input-Output Data, Trade Data, and Transaction Data

As mention in the previous section, there are numerous empirical studies examining the role of interfirm transactions in determining the spatial distribution of economic activities. These studies can be broadly divided into two groups depending on what type of transaction data are used, namely data from input-output tables (I-O tables) or international trade data. Starting with studies using I-O tables, Rosenthal and Strange (2001) and Ellison et al. (2010) use the aggregate volume of interfirm transactions by industry from I-O tables to investigate the relationship between the volume of transactions and the degree of agglomeration within an industry. Rosenthal and Strange (2001) found a positive correlation between the volume of transactions and the degree of agglomeration, while Ellison et al. (2010) showed that the volume of transactions between two industries has a positive effect on coagglomeration of these industries.

However, studies using I-O tables suffer from various limitations, since the data are aggregated at the industry level and there is no geographic information. For example, the data do not indicate whether geographically close firms have large volumes of transactions between them, and it is not clear whether proximity between firms promotes such transactions.

Turning to studies using international trade data, many of these estimate so-called gravity equations. International trade data consist of data on the value of trade transactions as well as geographic information on the transactions in the form of information on which countries were involved in those transactions. This information is used to estimate gravity equations (for examples, see McCallum 1995; Anderson and van Wincoop 2003, 2004). Such gravity equation estimations show that the volume of trade decreases with geographic distance. However, studies using gravity equations to examine interfirm transactions also suffer from shortcomings. For example, since geographical distance is typically measured as the distance between the capitals of two countries, such distance data are very imprecise in the case of countries with a large landmass. Further, trade data can only capture the transaction relationship between countries and is therefore useful only for understanding the country-level agglomeration of economic activities. However, most interfirm transactions actually take place within countries rather than across countries, and geographical frictions may differ in transactions within and between countries. Specifically, transactions tend to be heavily concentrated within an extremely close range and studies show that geographical friction increases non-linearly with distance. To consider the agglomeration of economic activities within a country, such as at the city-level, it is necessary to understanding geographic frictions within a country.

Given these shortcomings, recent years have seen an increase in the use of a third type of data, namely microdata on interfirm transactions, which identify the firms transacting with each other and provide much more detailed geographic information. Using transactions microdata makes it possible to conduct much more detailed analyses of interfirm transactions and the spatial distribution of economic activities than earlier studies using I-O tables. The following section provides a brief explanation of transactions microdata, while the section after that presents an overview of studies that have used such data.

5.3 What are Transactions Microdata?

Transactions microdata consist of data on interfirm transaction relationships. In Japan, for example, such data are mainly collected by private credit research firms. This is because when scrutinizing the credit information of a firm, all information about those with whom the firm deals is regarded as an important indication of that firm's creditworthiness. Typical and well-known examples of such data are those collected for the Teikoku Databank (TDB) and Tokyo Shoko Research (TSR) databases. The information in these databases is collected largely through interview surveys that ask questions about the main suppliers and customers of individual firms. The TSR dataset covers 826,169 firms, which is over half of all incorporated firms in Japan.

An example for the United States is the Commodity Flow Survey (CFS), which provides information on the shipper's business establishment and ship-to addresses at the zip code level. These data, supplemented with data on the transaction relationship between business establishments or firms, are also referred to as transactions microdata.

5.4 Empirical Research Using Transactions Microdata

This section provides a survey of studies that have examined the role of interfirm transactions as a determinant of agglomeration through the use of transactions microdata. As mentioned above, studies on the relationship between interfirm transactions and economic geography typically tended to rely on I-O tables or international trade data. The studies discussed here generally follow one of those two approaches, but introduce transactions microdata into the empirical framework.

Let us start by looking at studies that extend the use of I-O tables with transactions microdata. As mentioned above, the empirical approach in studies relying on I-O tables typically consisted of collecting data on the volume of trade within an industry from the I-O tables and then estimating the effect of the volume of trade on the degree of agglomeration of that industry. Using microdata on interfirm transactions makes it possible to examine transaction structures and their impact on industry agglomeration in much greater detail. An example is the study by Nakajima et al. (2013), which focuses not only on the volume of interfirm transactions within an industry, which can also be extracted from I-O tables, but also on the structure of the transaction network, which I-O tables cannot capture. Specifically, Nakajima et al. (2013) examine the inequality in the number of firms' transaction partners in intra-industry transactions using the Gini coefficient. The Gini coefficient can provide an intuitive grasp of whether there are hub firms in the network of transaction relationships within an industry. That is, a high degree of inequality indicates that a small number of firms function as hubs, so that transactions within the industry are highly concentrated.

The study by Nakajima et al. (2013) follows the empirical approach of Rosenthal and Strange (2001) to regress the agglomeration indicator proposed by Ellison and Glaeser (1997) on the intra-industry volume of transactions and the Gini coefficient on the number of transaction partners. As in previous studies using I-O tables, the coefficient on the intra-industry volume of transactions is positive and significant. In addition, the coefficient on the Gini coefficient for the number of transaction partners is negative and significant, and the result is robust even when controlling for other indicators of agglomeration determinants, such as knowledge spillovers and labor pooling. The result thus suggests that greater inequality in the number of business partners in a transaction network (that is, a greater role played by hubs) tends to be associated with a wider dispersion in the location of firms. A possible interpretation is that hub firms tend to purchase inputs of higher quality, so that transaction costs play a subordinate role and transaction partners hence are more dispersed. This finding shows that the structure of transaction networks, which cannot be identified in aggregate I-O tables, has a major impact on where firms in the network are located. This is a finding that could be obtained only through the use of transactions microdata.

Next, let us look at studies that extend the gravity approach. Hillberry and Hummels (2008) estimate gravity equations using domestic transactions data. Specifically, they employ microdata from the CFS and aggregate these data in very small geographical units (at the five digit zip code level). They find the following. First, the volume of transactions is concentrated in an extremely close range (at the five digit zip code level). Second, there are significant geographic frictions in inter-regional trade.

However, geographic friction in transactions may not necessarily be accurately estimated in the gravity equation approach. First, gravity equations require the aggregation of transaction microdata into appropriate spatial units. However, as highlighted by Hillberry and Hummels (2008), most transactions are over extremely short distances and are clustered within the same spatial unit (five digit zip code). This means that variation in transaction distances within the same spatial units disappears as the result of aggregating into spatial units. Second, as the number of spatial units increases, it becomes difficult in the gravity equation approach to control for the characteristics of economic activity in the origin and destination units. To control for the economic size and thickness in the origin and destination units, origin and destination fixed effects are often used (e.g., Harrigan 1996; Redding and Venables 2004). In this approach, an increase in the number of spatial units means that the number of variables increases linearly, making estimation computationally difficult. In the study by Hillberry and Hummels (2008), there are 29,194 spatial units and, given these challenges, they do not control for origin and destination effects.

To appropriately estimate geographical frictions, it is necessary to avoid aggregation and control for overall economic activity. One study along these lines is Nakajima et al. (2012b). Employing the K-density approach developed by Duranton and Overman (2005, 2008), the study empirically examines the agglomeration of transaction relationships between firms. The analysis starts by calculating all bilateral distances between firms that transact with each other and estimating the distribution of those distances. It then considers groups of firms that could potentially transact with each other and randomly assigns transaction relationships within the groups of firms that could potentially transact with each other. In the next step, the difference between the bilateral distance distribution between actual transaction partners and the counterfactual distribution that is generated from the random assignment of transaction partners is examined. Doing so makes it possible to measure the tendency of transactions to be concentrated within a relatively short range, while controlling for the tendency of firms to agglomerate. Nakajima et al. (2012b) results suggest that transaction relationships indeed are localized and tend to be concentrated within a range of 60 km. The results regarding the range of concentration is very similar to that obtained in Nakajima et al. (2012a), which focused on the concentration of firm locations. Taken together, the results of the two studies suggest that interfirm transactions are an important factor underlying industry agglomeration.

However, there are factors other than geographic distance, such as firm creditworthiness or size, which are extremely important when choosing a transaction partner. Nakajima et al. (2012a) do not consider the above factors in their analysis. Against this background, Nakajima (2012) extends the analysis using a structural estimation approach. Employing Fox's (2010) matching game estimation procedure, Nakajima (2012) formalizes firms' business partner selection problem and estimate firms' profit function. The study shows that even when other factors such as firm size and creditworthiness are controlled for, geographic distance has a negative impact on firms' profits. Further, the negative effect of distance on firms' profits is stronger in the case of sellers than buyers.

In sum, transaction microdata provide new insights on geographic frictions in transactions that cannot be detected using aggregate data.

5.5 Interfirm Transaction Networks and Corporate Decision Making

The preceding section focused on studies examining the geographic distribution of and frictions in interfirm transaction networks. Another strand of the literature on networks investigates the impact of networks on their members' decisions and behavior. Much of this type of research has concentrated on so-called "peer effects." Calvó-Armengol et al. (2009), for instance, examined the impact of friendship networks on the educational achievement of individual students. Such studies focus on human networks such as classmates, on neighborhoods, or on researcher collaborations. Extensive surveys of such studies have been conducted by Ioannides and Topa

(2010) and Durlauf (2004). The purpose here is to provide a brief overview of such research relevant to the study of interfirm networks rather than human networks.

There are a considerable number of studies examining peer effects in interfirm networks. Leary and Roberts (2014) showed that firms' decisions regarding corporate finances such their financial leverage are positively related to peer (competitor) firms' decisions. That is, an increase in peer firms' leverage ratio tends to be associated with increases in the own-firm leverage ratio. Fracassi (2012) found that corporate policy decisions depend on the decisions of peer firms measured in terms of the social ties of executives and directors of the firm. An increase in the social ties between two firms tends to be associated with an increase in the similarity of corporate decisions such as investment. Along similar lines, Bouwman (2011) found that networks of shared directors propagated firm governance policies to peer firms. Meanwhile, focusing on the networks of executives when they were MBA students, Shue (2013) defined peer networks by whether the executives were assigned to the same section at a business school and found that the firm outcomes were similar in peer firms. There are also a number of studies focusing on developing countries. Patnam (2012), for example, examined firms' peer networks measured in terms of interlocking board memberships in Indian corporations and found positive and significant peer effects with regard to corporate investment strategies and executive compensation. On the other hand, focusing on African countries, Fafchamps and Söderbom (2015) found limited network effects on business practices.

A considerable number of studies have examined the link between interfirm transaction networks and corporate decision-making. In particular, there has been considerable interest in the analysis of the foreign direct investment (FDI) decision of a firm. For instance, numerous studies have examined the impact of transactions between firms within affiliated corporate groups (*keiretsu*) in Japan on the choice of FDI location (Head et al. 1995, 1999; Belderbos and Carree 2002; Blonigen et al. 2005). In a recent study, Yamashita et al. (2014) used the TSR database to identify interfirm transaction networks and consider actual transaction relationships rather than *keiretsu* relationships. Specifically, they took into account the layered nature of transaction networks and considered the relationships of direct business partners (first tier), partners of these direct partners (second tier), and their partners (third tier). Based on this distinction, they found that only the location of first tier firms has a positive impact on the choice of firms' FDI location. That is, firms tend to choose a location where the firms' first tier transaction partners are located. Another study in this field is that by Itoh and Nakajima (2014) examining the impact that location in corporate transaction networks has on FDI decision-making. Their theoretical considerations suggest that centrality in a network—i.e., whether a firm is a hub firm in a network—is an important determinant of whether a firm conducts FDI. They follow this with an empirical analysis using Japanese firm-level FDI data, which supports their theoretical considerations.

5.6 Interfirm Transaction Relationships and the Geographic Propagation of Shocks

Another important aspect with regard to transaction relationship networks is the role they can play in spreading geographically localized shocks. A key example is the indirect damage to the economy caused by the disruption of supply chains following the Great East Japan Earthquake in March 2011. The supply chain disruptions not only forced firms in Japan reliant on suppliers from the affected region to halt operations, but had reverberations around the world. In other words, interfirm transaction relationships had the effect of widely propagating a geographically localized shock.

The example of the Great East Japan Earthquake illustrates that to understand the impact of natural disasters on the wider economy through interfirm transaction networks, it is important to study the structure of such networks and the geographic distribution of economic activities that explain how geographically localized shocks can spread across a nation or the globe. In fact, work in this important area of research has been growing in recent years. From a theoretical perspective, a study of particular interest is that by Henriet et al. (2011), which theoretically investigated the propagation of exogenous shocks such as a natural disaster through production-unit-level input-output networks. They found that redundancy of transaction partners and clustering of the production network (e.g., customers of a production unit are also suppliers to one of the suppliers of the production unit) have a positive effect on the economic robustness to shocks.

An example of an empirical study within this research strand is that by Saito (2013), which focuses on interfirm networks from a geographic viewpoint. Using TSR data, she examines the links between firms directly affected by the Great East Japan Earthquake and firms in the same network separated from them by three degrees (i.e., transaction partners of transaction partners of transaction partners of the affected firm). Her results show that although the share of firms that are directly affected by the Great East Japan Earthquake is extremely small, about 90 % of all firms in Japan are linked to an affected firm by three degrees of separation or less. This suggests that Japanese firms' transaction relationships have a so-called small-world structure, meaning that more or less all firms are linked with each other via a very small degree of separation. As a result, even geographically localized shocks that affect only a small number of firms tend to spread nationwide through interfirm transaction relationships.

As mentioned earlier, Nakajima et al. (2013) showed that industries in which transaction networks are characterized by the presence of hub firms tend be more geographically dispersed. Since a small-world structure primarily arises from the existence of hubs, it is possible that networks in industries characterized by the existence of hub firms have a small-world structure, meaning that due to the structure of such networks, geographically localized shocks spread relatively easily across the entire network. Given that such networks also tend to be more geographically dispersed, this suggests that such a network structure facilitates the nationwide spread of local shocks.

Finally, another closely related study is that by Todo et al. (2015), which examines how firm recovery from the Great East Japan Earthquake was affected by transaction networks. The results indicate that, controlling for other factors, firms with transaction partners outside the affected region were able to recover faster. Furthermore, regarding firms' medium-term recovery, they found that the number of transaction partners both inside and outside the affected region played an important role. This suggests cooperative relationships can facilitate the recovery of affected firms from a natural disaster in the medium run. However, in the short run, the number of transaction partners inside the affected area had a negative effect on firm recovery. This suggests that one of the costs of agglomeration is that the geographic concentration of transaction partners magnifies any direct or indirect damage caused by a disaster by eliminating transaction partners.

5.7 Concluding Remarks and Outlook for the Future

This chapter provided an overview of empirical research using transactions microdata to investigate interfirm transaction relationships and industry agglomeration as well as the spatial relationships between economic activities. It was suggested that microdata on interfirm transactions makes it possible to examine issues that could not be examined using input-output tables or trade data.

The overview concentrated on interfirm relationships based on physical transactions. However, relationships between firms are not necessarily limited to physical transactions. The theoretical studies by Borukhov and Hochman (1977), O'Hara (1977), and Fujita and Ogawa (1982) mentioned earlier, for instance, consider the communication between firms. Joint research conducted by firms may be said to be one form of such communication, since a joint research relationship can be interpreted as an exchange of implicit knowledge and ideas between firms, that is, a kind of knowledge spillover. Going back to Marshall (1890), such knowledge spillovers have been regarded as one determinant of agglomeration. Davis and Dingel (2012) theoretically show that such exchange of ideas between firms is one factor driving agglomeration in urban metropolitan areas. Inoue et al. (2013) studied geographic frictions in joint research relationships between establishments. They showed that joint research relationships between establishments for the creation of patents are geographically clustered. Further, joint research relationships between different firms are more geographically clustered. This suggests that geographical proximity facilitates collaboration between firms in knowledge creation.

Interfirm networks also include human networks between firms, such as worker flows. Recent studies using employer–employee data, such as Balsvik (2011), Parrotta and Pozzoli (2012), Stoyanov and Zubanov (2012) and Serafinelliy (2013) show that worker transfers between firms or plants have a positive impact on productivity through knowledge transfers.

As shown above, using transactions microdata in empirical research on interfirm networks can shed light on various unresolved issues in the field of spatial economics.

At the same time, such microdata also help to gain a better understanding of the structure of interfirm transaction networks. One challenge is to link—both from a theoretical and an empirical perspective—the analysis of complex network structures and the geographic distribution of economic activities. In the field of international economics, a number of scholars have attempted to model the complex structures of transaction networks (see, e.g., Costinot et al. 2013). Also of note is the study by Weisbuch and Battiston (2007), which provides a theoretical model of supply chain networks and the geographic distribution of economic activities.

The study of interfirm transaction networks and the spatial distribution of economic activities using microdata can be considered a frontier area of research from both a theoretical and empirical perspective. Research in this area therefore can be expected to yield many new, exciting findings.

References

Anderson, J., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *The American Economic Review, 93,* 170–192.

Anderson, J., & van Wincoop, E. (2004). Trade costs. *Journal of Economic Literature, 42,* 691–751.

Balsvik, R. (2011). Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *Review of Economics and Statistics, 93,* 285–297.

Belderbos, R., & Carree, M. (2002). The location of Japanese investments in China: Agglomeration effects, Keiretsu, and firm heterogeneity. *Journal of the Japanese and International Economies, 16,* 194–211.

Blonigen, B. A., Ellis, C. J., & Fausten, D. (2005). Industrial groupings and foreign direct investment. *Journal of International Economics, 65,* 75–91.

Borukhov, E., & Hochman, O. (1977). Optimum and market equilibrium in a model of a city without a predetermined center. *Environment and Planning A, 9,* 849–56.

Bouwman, C. (2011). Corporate governance propagation through overlapping directors, review of financial studies, forthcoming.

Calvó-Armengol, A., Patacchini, E., & Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies, 76,* 1239–1267.

Costinot, A., Vogel, J., & Wang, S. (2013). An elementary theory of global supply chains. *Review of Economic Studies, 80,* 109–144.

Davis, D., & Dingel, J. (2012). A spatial knowledge economy, NBER Working Paper #18188.

Duranton, G., & Overman, H. (2005). Testing for localization using microgeographic data. *Review of Economic Studies, 72,* 1077–1106.

Duranton, G., & Overman, H. (2008). Exploring the detailed location patterns of U.K. Manufacturing industries using microgeographic data. *Journal of Regional Science, 48,* 213–243.

Durlauf, S. (2004). Neighborhood effects. In V. Henderson & J. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, pp. 2173–2242). North-Holland: Amsterdam.

Ellison, G., & Glaeser, E. (1997). Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy, 105,* 889–927.

Ellison, G., Glaeser, E., & Kerr, W. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review, 105,* 889–927.

Fafchamps, M., & Söderbom, M. (2015). Network proximity and business practices in African manufacturing. *World Bank Economic Review* (Forthcoming).

Fox, J. (2010). Estimating matching games with transfers. mimeo.

Fracassi, C. (2012). Corporate finance policies and social networks. mimeo.

Fujita, M., & Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics, 12,* 161–196.

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: Cities, regions and international trade*. Cambridge: MIT Press.

Goyal, S. (2007). *Connections: An introduction to the economics of networks*. Princeton: Princeton University Press.

Harrigan, J. (1996). Openness to trade in manufactures in the OECD. *Journal of International Economics, 40,* 23–39.

Head, K., Ries, J., & Swenson, D. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International Economics, 38,* 223–247.

Head, K., Ries, J., & Swenson, D. (1999). Attracting foreign manufacturing: Investment promotion and agglomeration. *Regional Science and Urban Economics, 29,* 197–218.

Henriet, F., Hallegatte, S., & Tabourier, L. (2011). Firm-network characteristics and economic robustness to natural disasters. *Journal of Economic Dynamics and Control, 36*(1), 150–167.

Hillberry, R., & Hummels, D. (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review, 52,* 527–550.

Inoue, H., Nakajima, K., & Saito, Y. U. (2013). Localization of collaborations in knowledge creation, RIETI Discussion Paper Series, 13-E-70.

Ioannides, I., & Topa, G. (2010). Neighborhood effects: Accomplishments and looking beyond them. *Journal of Regional Science, 50,* 343–362.

Itoh, R., & Nakajima, K. (2014). Impact of supply chain network structure on FDI: Theory and evidence, RIETI Discussion Paper Series, 14-E-27.

Jackson, M. (2008). *Social and economic networks*. Princeton: Princeton University Press.

Krugman, P., & Venables, A. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics, 110,* 857–880.

Leary, M., & Roberts, M. (2014). Do peer firms affect corporate financial policy? *The Journal of Finance*, *69,* 139–178.

Marshall, A. (1890). *Principles of economics*. London: Macmillan.

McCallum, J. (1995). National borders matter: Canada-U.S. regional trade patterns. *The American Economic Review, 85,* 615–623.

Nakajima, K. (2012). Transactions as a source of agglomeration economies: Buyer-seller matching in the Japanese manufacturing industry, RIETI Discussion Paper Series, 12-E-21.

Nakajima, K., Saito, Y. U., & Uesugi, I. (2012a). Measuring economic localization: Evidence from Japanese firm-level data. *Journal of the Japanese and International Economies, 26,* 201–220.

Nakajima, K., Saito, Y. U., & Uesugi, I. (2012b). Localization of interfirm transaction relationships and industry agglomeration, RIETI Discussion Paper Series, 12-E-23.

Nakajima, K., Saito, Y. U., & Uesugi, I. (2013). Role of inter-firm transactions on industrial agglomeration: Evidence from Japanese firm-level data, RIETI Discussion Paper Series, 13-E-21.

O'hara, J. (1977). Location of firms within a aquare central business district. *Journal of Political Economy, 85,* 1189–1207.

Parrotta, P., & Possoli, D. (2012). The effect of learning by hiring on productivity. *RAND Journal of Economics, 43,* 167–185.

Patnam, M. (2011). Corporate networks and peer effects in firm policies. mimeo.

Redding, S., & Venables, A. (2004). Economic geography and international inequality. *Journal of International Economics, 62,* 53–82.

Rosenthal, S., & Strange, W. (2001). The determinants of agglomeration. *Journal of Urban Economics, 50,* 191–229.

Saito Y. (2013). The great east Japan earthquake and inter-firm geographical network, RIETI Discussion Paper, 12-J-020.

Serafinelliy, M. (2013). Good firms, worker flows and productivity. mimeo.

Shue, K. (2013). Executive networks and firm policies: Evidence from the random assignment of MBA peers. *Review of Financial Studies, 26,* 1401–1442.

Stoyanov, A., & Zubanov, N. (2012). Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics, 4,* 168–198.

Todo, Y., Nakajima, K., & Matous, P. (2015). How do supply chain networks affect resilience of firms to natural disasters? Evidence from the Great East Japan Earthquake. *Journal of Regional Science 55,* 209–229.

Weisbuch, G., & Battiston, S. (2007). From production networks to geographical economics. *Journal of Economic Behavior and Organization, 64,* 448–469.

Yamashita, N., Matsuura, T., & Nakajima, K. (2014). Agglomeration effects of inter-firm backward and forward linkages: Evidence from Japanese manufacturing investment in China. *Journal of the Japanese and International Economies, 34,* 24–41.

Kentaro Nakajima is an associate professor at the Graduate School of Economics, Tohoku University. Prior to his current position, he was formerly an associate professor, Endowed Chair in Regional Economy and Banking (The 77 Bank), Graduate School of Economics, Tohoku University (2008–2010) and an associate professor at the Institute of Economic Research, Hitotsubashi University (2010–2011). He earned his Ph.D. in Economics from the University of Tokyo in 2008. His research expertise covers spatial economics, economic development, and applied microeconometrics. He has published papers on topics such as localization of economic activities and productivity advantages in industrial clusters in Regional Science and Urban Economics, Journal of Regional Science, among others.

Chapter 6
Delineating Metropolitan Areas: Measuring Spatial Labour Market Networks Through Commuting Patterns

Gilles Duranton

Abstract This chapter first discusses the necessity of defining metropolitan areas and current practice in several countries. It argues for the use of a simple algorithm that exploits cross-municipality commuting patterns. Municipalities are aggregated iteratively provided they send a share of their commuters above a given threshold to the rest of a metropolitan area. This algorithm is implemented on Colombian data and its robustness is assessed. Finally, the properties of the resulting spatial labour market networks are explored.

Keywords Metropolitan area definition · Municipal aggregation · Colombian cities · Zipf's law

6.1 Introduction

This chapter proposes a methodology to define metropolitan areas by iterative aggregation of spatial units using daily commuting flows between them. In essence, a spatial unit A is aggregated to another spatial unit B if the share of the workers who work in B among all those that reside in A is above a given threshold. Another

---

Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, USA (e-mail: duranton@wharton.upenn.edu; website: https://real-estate.wharton.upenn.edu/profile/21470/). Also affiliated with the Centre for Economic Policy Research, the Rimini Centre for Economic Analysis, and the Spatial Economic Centre at the London School of Economics. I am grateful to Rafael Cubillos for giving me data without which this project would have been impossible. James Bernard, Matthew Degagne, and Hongmou Zhang provided very able research assistance. I also thank Paul Cheshire and Yoshi Kanemoto for very useful conversations and for getting me interested in this subject many years ago. Feedback from Alvaro Pachon, Rafael Cubillos, José Salazar and other seminar participants at Departamento Nacional de Planeación (DNP) is also gratefully acknowledged. This research was conducted when the author was working for the urban division of DNP. The collaboration of the entire division, Alejandro Bayonna, and Carolina Barco was greatly appreciated.

G. Duranton (✉)
University of Pennsylvania, 3620 LocustWalk, Philadelphia, PA 19104, USA
e-mail: duranton@wharton.upenn.edu

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_6

spatial unit C may next be aggregated to the union of A and B if, similarly, it sends a fraction of its commuters greater than the same threshold to this newly formed unit even though it may not have been possible to aggregate C directly to either A or B. This process of aggregation repeats until no further unit can be aggregated.

This algorithm is implemented with a threshold of 10 % of commuters on municipal data in Colombia to define metropolitan areas for this country, which currently lacks well-defined metropolitan areas. Although aggregating spatial units iteratively using a minimum commuting threshold is not novel, our implementation is novel in two respects. First, we show that a careful implementation of an aggregation algorithm that relies solely on a minimum commuting threshold criterion is enough to define meaningful metropolitan areas and generate metropolitan cores endogenously. This is unlike the practice of many statistical institutes. They usually predefine metropolitan cores and use a minimum commuting rule in conjunction with several other criteria. Second, we assess the robustness of the set of resulting metropolitan areas to changes in the minimum commuting threshold for aggregation.

Defining metropolitan areas is important for several reasons. Historically, as cities grew both in population and spatially, they would directly annex surrounding municipalities. In many countries, this process has stopped; richer municipalities resist fiscal integration with their poorer neighbours; mayors attempt to retain their jobs; or, as in Colombia, there may be significant constitutional and administrative barriers to merging municipalities. As a result, administratively defined cities are typically restricted to an urban core and are no longer representative of their broader metropolitan environment.

Related to this, existing administrative units such as municipalities do not generally constitute functionally autonomous units. Instead neighbouring municipalities are often economically integrated in all sorts of ways. This implies that an economic shock or a policy intervention in one municipality may have important spillover effects on its neighbours. Given the difficulty of keeping track of spillover effects, it is easier (and typically more efficient) for policies to target functionally consistent units.

Being able to deal with functionally consistent units is also important for research. For instance, cities tend to grow geographically by spreading outwards, outside the boundaries of the core municipality. When looking at patterns of urban growth based on municipal data, one may conclude that large cities grow slowly. This is often far from being the case. A core municipality is often 'full' and its metropolitan area typically grows at its extensive margin via its peripheral municipalities. Hence, urban growth is most appropriately measured at the metropolitan level.

Finally, cities constitute interesting spatial networks of commuting workers, transacting firms, or interacting individuals. To be able to study these networks meaningfully it is fundamental to be able to describe them first.

The rest of this chapter is organised as follows. Section 6.2 provides some background about the situation in Colombia, current practice in other countries, and prior academic literature. Section 6.3 presents the data and our aggregation algorithm. Section 6.4 provides our list of metropolitan areas and metropolitan regions for Colombia. The robustness of these results is assessed in Section 6.5. Finally, Section 6.6 concludes.

6.2 Background, Current Practice, and Literature

6.2.1 The Current Situation in Colombia

Although there exists an official set of 'metropolitan areas' in Colombia, these areas are mostly administrative units, constituted on a voluntary basis (Senado de la República 2012). While there is certainly a strong case for associations of neighbouring municipalities to form broader formal institutions, these 'political' metropolitan areas are usually not appropriate for analysis and decision-making by higher levels of governments.[1]

For historical reasons and, perhaps, because of institutional rivalry, Bogotá, the largest municipality in Colombia, is not part of any metropolitan area even though there is no observable discontinuity between Bogotá and, for instance, its Southern neighbour Soacha. Cali, the third largest city in Colombia, is not part of any institutional arrangement with any of its neighbours either. Barranquilla is a less extreme case. Its official metropolitan area is composed of only five municipalities whereas we obtain a metropolitan area of already nine municipalities with an extremely conservative commuting threshold of 30 % (which is three times as large as our preferred threshold of 10 %). On the other hand, Medellín, the second largest city, has formed the 'metropolitan area of the Aburrá Valley' which corresponds exactly to the one generated by our algorithm with our preferred commuting threshold of 10 %. However, Medellín is the exception, not the rule. A systematic and consistent set of metropolitan areas is needed for Colombia.

6.2.2 Current Practice in the World

While details vary, there are two features that are common to most ordinarily used definitions of metropolitan areas.

The first is the preponderant role given to commuting patterns. Metropolitan areas are thus viewed as integrated labour markets. There are good reasons for this. Since Marshall (1890), economists usually think of cities as bringing benefits, in terms of 'thick' labour markets, greater diversity of available final and intermediate goods, and more intense individual interactions conducive to knowledge spillovers. Focusing on the first series of these benefits coming from local labour markets makes sense for two reasons. The first is that commuting patterns can easily be tracked. The census and many other sources of labour market data usually record both the place of residence and the place of work of workers. The variety of final and intermediate goods,

[1] The French government defines 'statistical' metropolitan areas though its statistical institute (INSEE). At the same time, there are many 'urban communities' which are voluntary unions of neighbouring municipalities, i.e., political metropolitan areas. The two differ, sometimes considerably, but coexist to serve extremely different purposes.

input-output linkages, and knowledge spillovers are much more complicated to track (Holmes 1999; Charlot and Duranton 2004; Handbury and Weinstein 2010). There is also a broad consensus among economists interested in cities that commuting patterns usually take place over distances that we naturally recognise as being 'metropolitan'. Instead, knowledge spillovers might take place over much shorter distances while input-output links often take place at a scale broader than the metropolitan area, as argued for instance by Krugman (1991).

In addition, there are other criteria that could be used to define metropolitan areas including non-economic criteria such as the sense of belonging to a place, etc. In practice, however, because they are easier to track and because their scale seems right, commuting patterns play an overwhelming role in the definition of metropolitan areas.

The second key feature of most official definitions of metropolitan areas is the use of an iterative approach to aggregate municipalities (or other basic geographical units such as counties in the US) into metropolitan areas. More specifically, a minimal threshold of commuters is chosen. As soon as the share of commuting flows from an origin municipality to a destination municipality is above this threshold, the origin municipality is aggregated to the destination municipality. We will refer to the aggregated municipality as a 'satellite' municipality and the one it is aggregated to as its 'core'. These two municipalities become part of the same metropolitan area. This procedure is then repeated until there remains no municipality to aggregate.

If employment in metropolitan areas were fully centralised at a unique central business district, there would be no need to use an iterative approach. All relevant municipalities would be aggregated in a single round of aggregation. However, in reality only a small proportion of jobs is concentrated in the centre of metropolitan areas. Glaeser and Kahn (2001) argue that less than 10 % of employment in US cities is concentrated within 5 km of their centre. This is far from the idealised description of monocentric cities where all the jobs are located in a central business district (Alonso 1964; Muth 1969; Mills 1967). As a result, and given the gravitational nature of commuting where the number of commutes decreases with distance an iterative aggregation procedure is needed. Imagine a core municipality A, a 'first-ring' municipality B, and a 'second-ring' municipality C. Municipality C may be sending lots of commuters to A and B but not enough to warrant immediate aggregation to A. As a result, B may be aggregated to A at the first round. Then C will be aggregated to the union of A and B at the second round.

Note that commuting thresholds are defined relative to the number of workers in the municipality at hand. This is because municipalities differ vastly in terms of their resident labour force. Using a relative threshold is important because it allows the aggregation of a small municipality that sends all its residents to the core municipality. Using an absolute threshold would not allow for this. Worse, on Colombian data it would lead to very misleading outcomes since there are many 'commuters' (in absolute terms) between the largest cities, including for instance the pair composed of Bogotá and Barranquilla, which are located several 100 km from each other. Looking at absolute numbers of commuters is an interesting measure of 'links' between municipalities and could perhaps be instrumental in the circulation of knowledge. This, however, does not help aggregate nearby municipalities into metropolitan areas.

Aside from these two features which are used by most countries that define metropolitan areas, there are several others which are common to many countries.

The first of these features (which is used for instance in the US) is the pre-determination of a 'core'. That is, the authority in charge of defining metropolitan areas will aggregate satellite units (counties in the American case) only around particular 'core' units which satisfy *ex ante* some particular properties in terms of population size and density. Put differently, a city needs to be 'big enough' and 'dense enough' to be considered as a potential nucleus for a metropolitan area. For instance, in the US, the core county must "(a) Have at least 50 % of [its] population in urban areas of at least 10,000 population; or (b) Have within [its] boundaries a population of at least 5000 located in a single urban area of at least 10,000 population." (US Office of Management and Budget 2010).

While this type of criterion seems intuitive, our results for the Colombian case show that it is not needed in practice. First, pre-defined cores might be arbitrary. Instead the algorithm that defines metropolitan areas should also pick the cores endogenously. Then, given the absence of *ex ante* cores, issues surrounding the minimum criteria that a core should satisfy become moot, which is desirable. As will become clear below, it is best to avoid criteria that are either un-necessary or that can be manipulated to define metropolitan areas whimsically.

It is possible to think of some mostly rural municipalities that would attract a significant fraction of commuters from other larger municipalities. These rural municipalities would then be perversely tagged as 'metropolitan cores'. One could also imagine large groups of rural municipalities with lots of cross-commuting giving rise to 'metropolitan areas'. They would obviously be missing 'urban character'. While such pathological situations are theoretically possible in the absence of pre-defined cores, the Colombian example shows that in all cases aggregation into metropolitan areas occurs around the largest municipality and there are very few cases of aggregation involving municipal cores with a small population. As argued below, these areas can always be selected out *ex post* by imposing a minimum population size for metropolitan areas.

Geographical contiguity may also be added as a criterion to define metropolitan areas. This seems natural. A highly integrated area is expected to be geographically continuous (also sometimes referred to as coterminous). While there might be esthetic reasons for imposing geographical continuity, there is no strong economic reason. Two municipalities separated by inhospitable terrain may form one economically integrated area and the area in-between may remain mostly rural. It is not clear why this area in-between should be forcibly integrated when it is not interacting with the other two municipalities. In any case, this is again a moot point because the algorithm used below to define metropolitan areas only aggregates contiguous areas with our preferred threshold of 10 %. Again the gravitational nature of commuting implies that a municipality completely surrounded by a metropolitan area is unlikely to remain alone when all of its neighbours have been aggregated. In any case, rather than impose a contiguity constraint *ex ante* it is better to check for exceptions *ex post* and attempt to understand them.

Statistical authorities also sometimes add further criteria, including asking for 'local opinions' in the US. A related issue is whether the algorithm used to define metropolitan areas should be applied in a strict fashion or instead be used more 'flexibly'. Conceptually, these two questions are separate. One may want to use a complicated algorithm to define metropolitan areas and apply it in a strict manner. Alternatively, it is possible to think of a simple algorithm subject to some 'operational adjustments' *ex post*. In practice, the issues of the number of criteria in the algorithm used to define metropolitan areas and whether this algorithm is applied flexibly or not are deeply intertwined. The use of many criteria (including fairly subjective ones that rely on local opinions) is probably a way to have some flexibility in the delineation of metropolitan areas. To make things worse, countries that use many criteria do not make their exact algorithm, the inputs into it, and its output public.

There are two reasons why one should use a simple and transparent algorithm that is applied strictly to define metropolitan areas. The first is that it really makes no sense to develop a methodology that defines metropolitan areas if it is to be renegotiated *ex post* because of a statistician's whims or because of political pressure. The second reason is that metropolitan areas are part of economic policies in some countries. Hence, the delineation of metropolitan areas affects the allocation of resources. It then becomes easy to see how and why the definition of metropolitan areas can become politicised. Policies that allocate resources to metropolitan areas whose definitions have been meddled with are biased. This means outcomes that are poorly measured, less efficient, and potentially unfair. To avoid political interference, it is crucial that the definition of metropolitan areas remains as simple as possible and that the task of defining them be given to an independent statistical institute. The advantages of doing so are overwhelming relative to the possibility of having one or two 'awkward' cases in the final list of metropolitan areas (beyond a list of metropolitan areas).

Statistical institutes also sometimes impose an *ex post* minimum population size criterion for metropolitan areas. This may not be needed if, for instance, the definition of metropolitan areas already imposes some minimum population constraint for the core municipality. For policy purpose, it is obvious that a minimum size threshold often needs to be considered. This threshold may depend on the type of policy. Looking at the provision of university education for which the metropolitan area is arguably the relevant spatial scale, it is clear that a relatively high population threshold needs to be considered. We cannot expect 'metropolitan areas' with a few thousand inhabitants to be provided with universities. When looking at environmental issues such as the disposal of solid residuals, it is probably best to consider all metropolitan areas including small 'lone' municipalities. It is also the case that imposing a stringent threshold on the entire list of metropolitan areas suppresses useful information. As a result it is usually best to generate a complete list of municipalities and metropolitan areas. A cutoff can then be imposed for a particular analysis or a specific policy or set of policies. This has the added benefit of allowing for more relevant cutoffs to be considered and forcing the policy makers to justify their chosen threshold in a clear and transparent manner.

Another interesting feature of the definition of metropolitan areas in many countries is the fact that there are often several such definitions. For instance, France defines both ‘urban areas’ and ‘urban units’. The latter are typically organised around a single core whereas the former are more standard (and broadly defined) metropolitan areas. The same situation is encountered in the US where there is a list of ‘consolidated’ metropolitan areas and a list of ‘primary’ metropolitan areas. Consolidated metropolitan areas are the union of several adjacent primary metropolitan areas. To be more concrete, Washington DC and Baltimore form two separate metropolitan areas but they also belong to the same consolidated metropolitan area. Again, like in France, the primary metropolitan areas appear to be core-based and correspond to integrated labour markets. Instead, consolidated metropolitan areas capture broader spatial units, and perhaps other forms of economic integration. Baltimore and Washington DC are certainly part of the same ‘economic region’ even though the proportion of workers that commute to DC from the Northern suburbs of Baltimore is probably quite low. There is a clear tradeoff here. Having multiple definitions will allow policy makers and analysts to capture different dimensions of economically integrated areas. At the same time, a multiplicity of definitions opens the door to arbitrary decisions and political interference. There is also the issue of how to go on about several definitions and whether they should be based on different thresholds for commuting or defined by different principles. While we return to these issues in the Colombian case below, it is hard to deny that having two different definitions for two different spatial scales is attractive.

We draw the following conclusions from the preceding discussion. The case for defining metropolitan areas based on commuting flows and for using an iterative procedure is extremely strong. The case for having two definitions to capture two different scales is also strong. On the other hand, the justification for many other practices routinely used by statistical institutes appears weak. Defining ‘cores’ ex-ante seems un-necessary, prevents useful checks on the algorithm, and opens the door to political interference. The same arguments apply with respect to the use of other (i.e., non-commuting) criteria to define metropolitan areas. Finally, using a simple and transparent algorithm that can be replicated (or used by others) allows for a number of useful checks. The usual practice of statistical institutes of proposing ‘a list’ of metropolitan areas without the raw data and the details of their algorithm is clearly unsatisfactory.

6.2.3 Prior Literature

The necessity to define urban areas first became clear in the US during the 1950s. Strong urban expansion and suburbanisation was no longer accompanied by municipal annexation. This led to a divergence between the political boundaries of the urban core and the economic boundaries of metropolitan areas. To resolve this problem, the US bureau of the census defined its Standard Metropolitan Statistical Areas (SMSAs) in the early 1950s. Early discussions in Berry (1960) and Fox and Kumar

(1965) were very much focused on defining metropolitan areas within a central place theory framework. Later Berry et al. (1969) offered a remarkable early discussion that echoes many of the points made here and suggested relying solely on commuting patterns towards a predetermined urban core to define metropolitan areas. Following the US, other developed countries also started defining their own metropolitan areas without seemingly much academic input. Their choices came under scrutiny in Hall and Hay (1980) and Cheshire and Hay (1989) who attempted to develop a broader perspective on European cities and needed a consistent set of units.

More recently, Kanemoto and Kurima (2005) have proposed an algorithm for Japan that has been widely used by subsequent research in absence of an official definition for metropolitan areas for this country. There is also a small stream of research that assesses how a range of local economic outcomes spatially autocorrelates across small spatial units to aggregate them into larger ones (see Cörvers et al. 2009, for a recent example). In this spirit, a particularly interesting variable is used by Bode (2008): land prices. He first detects some centres. These centres are defined as spikes of land prices that are statistically significant. He then estimates the part of urban land prices at each location that can be attributed to these centres and aggregates satellite areas accordingly. His approach is interesting as land prices are expected to reflect many different types of interactions across places beyond commuting. The main drawback is that a lot of structure is imposed and the results may be sensitive to minor aspects of this structure. Finally, geographers often propose lists of metropolitan areas but the definitions they propose are usually ad-hoc (see for instance Molina 2001, for Colombia).

We also note that extant research sometimes defines its own zoning (Briant et al. 2010; Rozenfeld et al. 2011). The delineations currently used by researchers differ a lot. Using different zonings for policy purposes may be an issue because it is well known that the zoning that one adopts may drive some of the results.[2] At the same time, as already argued, there is nothing intrinsically wrong with using different zonings for different purposes since different problems may require a focus on different spatial scales. There is also a strand of literature (e.g. Duranton and Overman 2005) that attempts to measure economic phenomena in continuous space doing away with spatial units altogether. This is not an option here given our perspective.[3]

6.3 A Simple Aggregation Algorithm

In agreement with the argument above, our proposed algorithm is as simple as possible. It aggregates a spatial unit to another if the former sends a high enough fraction of its commuters to the latter. Subsequently, a third spatial unit is aggregated to the union of the first two provided it sends a high enough fraction of its commuters to

---

[2] See for instance the well known 'Modifiable Areal Unit Problem' (MAUP). See Cressie (1993) for a presentation and a discussion.

[3] For instance, it is obvious that policies that allocate money to 'places' need discrete spatial units.

this newly formed unit even though it may not have been possible to aggregate this third spatial unit to any of the first two individually. This process is repeated until no spatial unit remains to be aggregated.

6.3.1 Preliminary Issues

Before going deeper into the details of the algorithm and its implementation to the Colombian case, it is useful to discuss the choice of commuting threshold. This choice is fundamentally arbitrary. Theory offers no reliable guidance here because economic integration between places follows a continuum. But since the objective is to delineate discrete units, there is no way around the necessity of a threshold. Taking a high threshold leads to the aggregation of very few satellite municipalities to urban cores, whereas taking a low threshold will lead to extremely large metropolitan areas. At the extreme, if each municipality sends at least one commuter to each of its neighbours, taking an arbitrarily low threshold will imply only one metropolitan area that covers the entire country. This is not helpful.

Adding to this, the choice of threshold is likely to depend on the size of the underlying units to be aggregated. Colombian municipalities are fairly large (on average more than 100 km^2). The gravitational nature of commuting implies that large municipalities will send on average only a small proportion of their commuters to work in other municipalities. Instead, France has more than 35,000 municipalities (and their average land area is only about 15 km^2). We thus expect much higher commuting flows between French municipalities because of this. Unsurprisingly, the threshold used by the French statistical institute is high at 40 %.

Commuting distances also depend on the level of development. In developed countries where a large fraction of workers can commute by car or using well-developed public transportation systems, a large proportion of workers may be able to commute over long distances. In Colombia, where car ownership is still limited and public transportation underdeveloped, the fraction of commuters that can commute over long distance is much lower than in Europe or North America. Hence, one may want to use different thresholds in developed and developing countries. This said, we also need to keep in mind that it is desirable to retain some consistency for the definition of metropolitan areas as a country develops.

A related problem associated with the choice of commuting threshold is the sensitivity of the delineation of metropolitan areas to small changes in the threshold. This can occur because of the iterative nature of the algorithm. Think of the following hypothetical example. Municipality D sends 12 % of its workers to municipality C and 10 % to B. Municipality C sends 12 % of its workers to B and 10 % to A. Finally, municipality B send 19 % of its workers to A. With a commuting threshold of 20 %, all four municipalities remain isolated since there is no flow above this threshold. For a threshold below 19 %, B gets aggregated to A at the first round. Then C which sends $10\% + 12\% = 22\%$ to the union of A and B get aggregated at the second round. At the third round, D also get aggregated so that we end up with a metropolitan area made of all four municipalities. In this example, a small change of threshold from 20 % to below 19 % leads to a radically different zoning.

To the possibility of such perverse cases suggested by this example, there are two responses. The first is to choose a 'natural' threshold (typically a round number) to avoid any suspicion of interference. The second response is to assess the sensitivity of the delineation of metropolitan areas with respect to the choice of threshold by comparing outcomes for different values of the commuting threshold. We perform such robustness checks below.

6.3.2 Data

To define metropolitan areas for Colombia, we use the matrix of commuting flows from the 2005 Colombian census and 2010 population estimates from the Colombian statistical institute, DANE. This choice reflects two conflicting constraints: using consistent data (preferably from the same year) and using the most recent data. For population, the entire population of each municipality was considered. For each municipality, Colombian statistics typically distinguish between an urban (or 'head') part and a rural part. Taking the entire population has the obvious drawback of aggregating rural populations to metropolitan areas. This drawback is minor in practice since most of the population of the municipalities that form large metropolitan areas is overwhelmingly 'urban'. Discarding rural populations would also lead to some awkward choices to be made about how to compute commuting shares since data for commuting flows are only available for entire municipalities.

Census populations or population estimates based on censuses are the best available population data in most countries including Colombia. Commuting flows are measured from only a subsample of the population surveyed by the Colombian census.

This follows common practice in most countries where commuting questions (together with lots of other questions) are usually administered through the 'long forms' of the census given only to a fraction of the population for cost reasons. In our case, this suggests some minor imprecisions due to mismeasured commuting flows. The lack of precision becomes more important as lower commuting thresholds are considered since with a low threshold of say 1 %, we may be well below the statistical margin of error in smaller municipalities. Results for low threshold are reported below but some care is needed in their interpretation given this reliability issue.

To delineate metropolitan areas for Colombia, we propose a commuting threshold of 10 % which appears reasonable given that Colombian municipalities are fairly large.

6.3.3 Algorithm

The algorithm is available upon request. It was programmed in Stata. After cleaning up the original matrix of cross-municipality commuting flows and creating a number of working files, each loop of aggregation works as follows. Among all pairs of origin and destination municipalities the algorithm flags those for which the share of commuters from the origin is above the chosen commuting threshold. Before being aggregated to a destination, the algorithm verifies that in case a municipality could be aggregated to several destinations, it is uniquely aggregated to the one it sends the most workers to. In case commuting flows between two municipalities are above the threshold in both directions, the algorithm also makes sure that the smallest municipality is aggregated to the largest.

At the aggregation stage, the name of the origin municipality is appended behind the name of the destination municipality and populations are added. The matrix of commuting flows is also appropriately aggregated and redefined. For instance, if municipality C sends 8 % of its workers to municipality B and 9 % to municipality A and if B is appended to A, then the commuting flows from C to B and C to A are aggregated into a unique flow of 17 % from municipality C into the metropolitan area A + B. The process of selection of commuting flows and aggregation is then repeated until no municipality or group of municipalities remains to be aggregated to a metropolitan area.

As final output, the algorithm produces a list of metropolitan areas with its component municipalities (a 'core' and its 'satellites') and single municipalities. For verification purposes, the algorithm also keeps track of all origin municipalities which were aggregated and the destination municipalities they were aggregated to.

This algorithm generates a list of metropolitan areas and municipalities associated with a given commuting threshold.

We also propose to define broader units, urban regions. As argued above this is in-keeping with existing practice in many countries. Recall that, for instance, the US metropolitan areas of Washington DC and Baltimore are separate but they are also part of the same 'consolidated' metropolitan area. To define these broader urban regions, a natural approach would be to use the same principle as with metropolitan areas but adopt a lower commuting threshold. For these urban regions, we take a lower threshold of 5 % but note that this change alone does not lead to dramatically larger units and clearly falls short of the notion of 'broad urban region'. The natural temptation would then be to lower this threshold even further. This is not a good idea since, as argued above, the aggregation exercise becomes fragile with very low commuting thresholds.

There is a deeper reason why even extremely low aggregation thresholds do not lead to broad urban regions. This is due to the self-reinforcing nature of the iterative aggregation process used to delineate metropolitan areas. To understand this subtle point, it is best to take a concrete example from Colombia. The 'coffee region' of Colombia is a confined high plateau between two branches of the Andes. It has three major cities which are fairly close to each other. The municipality of Pereira has around 450,000 people, that of Manizales is slightly below 400,000, and that of Armenia is slightly below 300,000. As small neighbouring satellite municipalities get aggregated to these three core municipalities, the three metropolitan areas that they form get more ‘entrenched’. The municipalities that are in-between these three main cities may see a fair amount of cross-commuting. But, as aggregation proceeds, these ‘in-between’ municipalities get aggregated to one of the three cores together with more peripheral municipalities. Given the gravitational nature of commuting, the aggregation of these peripheral municipalities lowers the tendency for their metropolitan areas to commute with each other. As a result, these metropolitan areas do not merge into a large single urban region even for a commuting threshold as low as 1 %.[4] However, it is interesting to observe that in many cases we obtain metropolitan areas that are contiguous with each other. Hence to define metropolitan regions, we propose to aggregate metropolitan areas that are contiguous with each other with a commuting threshold of 5 %. Then, the three separate areas aggregated around Pereira, Manizales, and Armenia, which are contiguous, are also the main centres of the larger urban region of Pereira-Manizales-Armenia.

6.4 Results

6.4.1 Metropolitan Areas

For the preferred commuting threshold of 10 %, the list of the 45 resulting metropolitan areas with more than 100,000 inhabitants is provided in Table 6.1. There are another 39 metropolitan areas with population above 50,000. In total, 99 satellite municipalities get aggregated to 22 cores, 19 of which have a population above 100,000. All the other municipalities remain stand alone municipalities. Metropolitan areas with a population above 100,000 are also depicted on the map of Fig. 6.1.

Before going further into the description of the list of metropolitan areas in Table 6.1, a few important features related to the algorithm need to be discussed. First, the iterative nature of the algorithm is fundamental. With a 10 % threshold, the algorithm goes through seven rounds of aggregations before converging. In the case of the largest metropolitan area composed of Bogotá and 22 neighbouring satellite municipalities, only nine of them are added at the first round of aggregation.

It is also interesting to note that the algorithm always picks as core municipality the largest municipality of the metropolitan area. This demonstrates that defining cores *ex ante* is unnecessary in practice. As can be verified on the map of Fig. 6.1, the metropolitan areas generated by the algorithm are also composed of contiguous municipalities. This shows that imposing contiguity is not needed either. Finally, there is no set of small and rural municipalities that get aggregated into much larger

[4] This phenomenon is not unique to the coffee region. The same is observed in the region of the Caribbean coast where three of the main cities: Barranquilla, Cartagena and Santa Marta do not merge even for a low commuting threshold of 1 %.

'metropolitan' areas. It is clear from the list given in Table 6.1 that the aggregation of peripheral municipalities into broader metropolitan units occurs mostly for the largest municipalities.

Table 6.1 Colombian metropolitan areas with population above 100,000. (Sources: DANE and author's computations)

Core municipality	Metropolitan municipalities	Metropolitan population	Core population	Large satellite municipalities
Bogotá DC	23	8,672,087	7,363,782	Cajicá, Chía, Funza, Mosquera, Soacha, Facatativá, Madrid, Zipaquirá
Medellín	10	3,544,703	2,343,049	Bello, Caldas, Copacabana, Envigado, Itagui, La Estrella
Cali	10	2,719,683	2,244,639	Candelaria, Jamundí, Yumbo, Florida, Pradera
Barranquilla	16	2,214,344	1,186,640	Baranoa, Malambo, Sabanalarga, Soledad
Cartagena	7	1,142,697	944,250	Arjona, Turbaco
Bucaramanga	4	1,074,929	524,112	Floridablanca, Girón, Piedecuesta
Cúcuta	4	773,659	618,310	Los Patios, Villa del Rosario
Pereira	3	717,383	457,103	Dosquebradas
Ibagué	1	526,547	526,547	
Santa Marta	1	447,857	447,857	
Villavicencio	2	441,906	431,476	
Manizales	2	439,630	388,525	
Armenia	4	430,749	288,908	Calarca
Pasto	2	416,224	411,706	
Montería	1	409,476	409,476	
Valledupar	1	403,414	403,414	
Buenaventura	1	362,625	362,625	
Neiva	1	330,487	330,487	
Rionegro	5	296,614	110,329	
Palmira	1	294,580	294,580	
Popayán	1	265,702	265,702	
Sincelejo	1	256,241	256,241	
Tuluá	2	217,189	199,244	
Riohacha	1	213,046	213,046	

Table 6.1 (continued)

Core municipality	Metropolitan municipalities	Metropolitan population	Core population	Large satellite municipalities
Tunja	5	200,696	171,082	
Barrancabermeja	1	191,498	191,498	
San Andres de Tumaco	1	179,005	179,005	
Sogamoso	9	165,183	115,564	
Florencia	1	157,450	157,450	
Apartadó	1	153,319	153,319	
Uribia	1	144,990	144,990	
Maicao	1	141,917	141,917	
Turbo	1	139,628	139,628	
Girardot	3	139,155	101,792	
Ipiales	2	129,808	123,341	
Cartago	1	128,566	128,566	
Yopal	1	123,361	123,361	
Magangué	1	122,913	122,913	
Fusagasugá	1	121,535	121,535	
Guadalajara de Buga	1	116,105	116,105	
Quibdó	1	114,548	114,548	
Duitama	2	114,470	110,418	
Lorica	1	114,145	114,145	
Pitalito	1	113,980	113,980	
Ciénaga	1	103,066	103,066	

Name of satellite municipalities reported above a population of 50,000

The list of the 84 largest metropolitan areas contains 180 municipalities (of more than 1100 for the entire country). These 84 metropolitan areas host 32.1 million people, or about 71 % of the population of Colombia. We note that peripheral municipalities are concentrated around the largest four cities. 55 of the 99 satellite municipalities are aggregated to one of the four largest Colombian municipalities. We also note that only four satellite municipalities are aggregated to core municipalities to form metropolitan areas with a population below 50,000 inhabitants. There is a strong rank correlation between the ranking of metropolitan areas in terms of population and the corresponding ranking of their core municipalities. For metropolitan areas with a population above 100,000, the correlation of log of population between the metropolitan area and the core municipality is 0.98. This said, there is some variation. The municipality of Medellín, the second largest in the country, has a

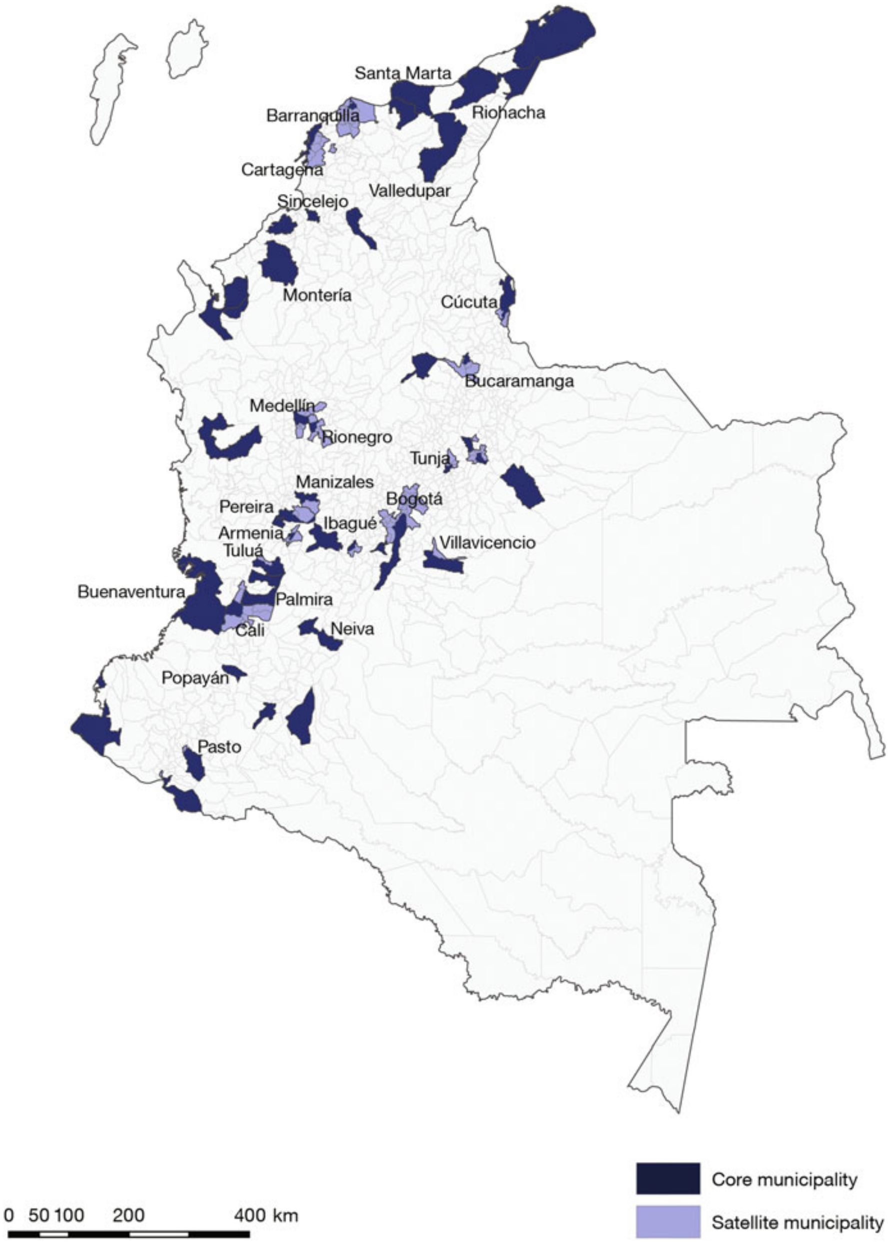


Fig. 6.1 Map of Colombian metropolitan areas with population above 100,000. (Sources: DANE and author's computations. Notes: Name of metropolitan area reported above a population of 200,000)

population only 4 % larger than that of the municipality of Cali, the third largest. However the metropolitan area of Medellín has a population 30 % larger than that of metropolitan Cali.

Viewed differently, our aggregation into metropolitan areas corrects for the idiosyncracies of the delineation of Colombian municipalities. The municipality of Medellín is geographically relatively small whereas that of Cali is large. At one extreme, in the cases of Barranquilla or Bucaramanga, the metropolitan area has a population that is twice that of the core municipality. At the other extreme, some large municipalities like Santa Marta, Ibagué, or Villavicencio either remain isolated or only receive tiny satellite municipalities so that their metropolitan population roughly coincides with their municipal population. The near absence of satellite for these municipalities is unsurprising. Santa Marta is a coastal city in decline and residents of neighbouring municipalities will be more easily lured to work in Barranquilla which is located fairly close. Ibagué and Villavicencio are fairly large isolated cities located close to major geographical 'ruptures'.

The four panels of Fig. 6.2 provide four magnified maps of the four most important concentrations of urban population where 16 of the largest 20 metropolitan areas are located including the largest five. These maps illustrate cases of contiguous metropolitan area such as Medellín and neighbouring Rionegro or the main cities of the coffee regions. These cases suggest that it is indeed interesting to consider a regional level of aggregation above metropolitan areas.

Overall the output generated by the algorithm appears to be highly consistent with both the underlying principles exposed above and qualitative features of the urban geography of Colombia.

6.4.2 Urban Regions

We now turn to the delineation of broader urban regions. To define these regions we take a lower commuting threshold of 5 % and aggregate the resulting metropolitan areas that are adjacent into urban regions.

The list of the 27 resulting urban regions composed of at least one metropolitan area of more 100,000 inhabitants is provided in Table 6.2. These urban regions are also depicted on the map of Fig. 6.3, panel (a).

Several features stand out from Table 6.2 and from the map of Fig. 6.3 a. The most important is the emergence of several important urban regions composed of a number of metropolitan areas. The Caribbean coast along the Cartagena–Santa Martha axis appears as the second most important urban region of the country with more than 4 million inhabitants.[5] There is also significant consolidation around Cali,

[5] This region is technically contiguous with the Valledupar-La Guajira region to its north-east. However, this contiguity is minimal and the Sierra Nevada mountain separates these two regions which are probably best treated as separate. Going from Santa Marta to 'neighbouring' Valledupar is a 5 h drive. Should these two regions be treated as one, they would form a region with 5.3 million inhabitants over 50 municipalities.

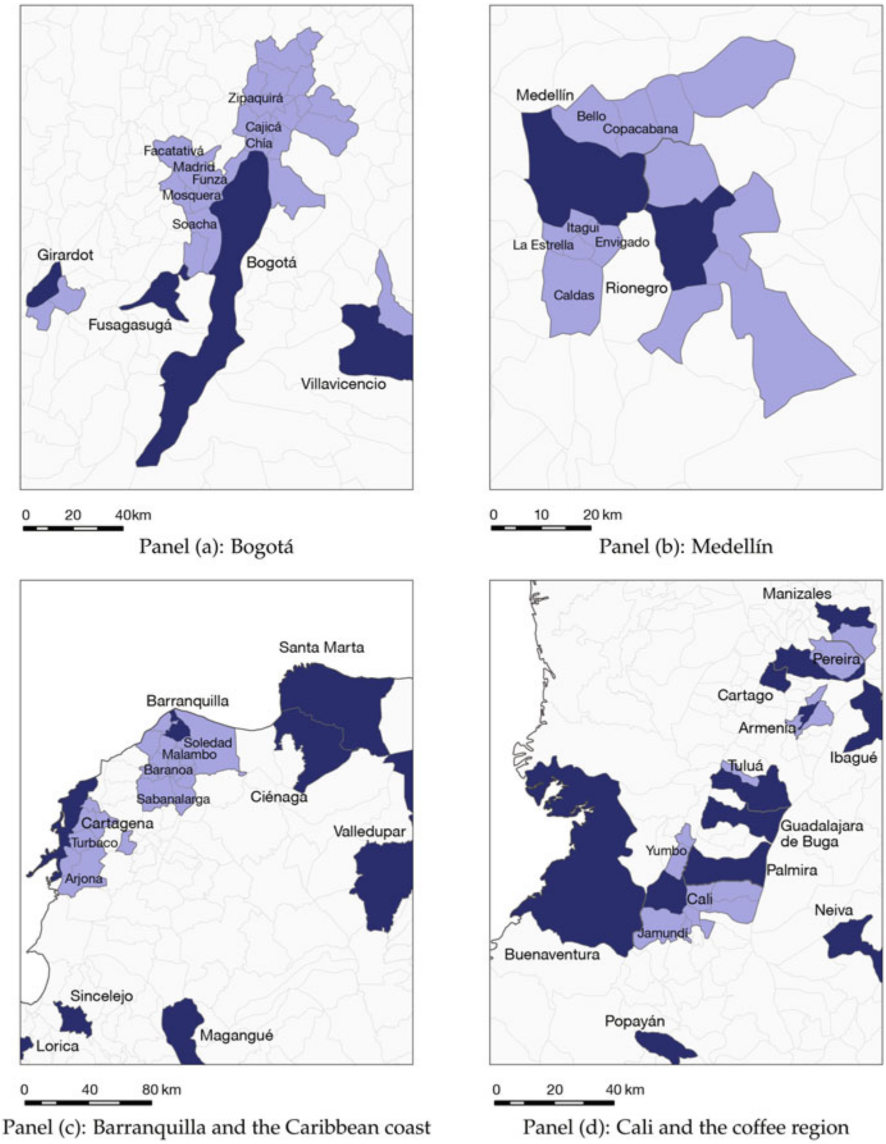
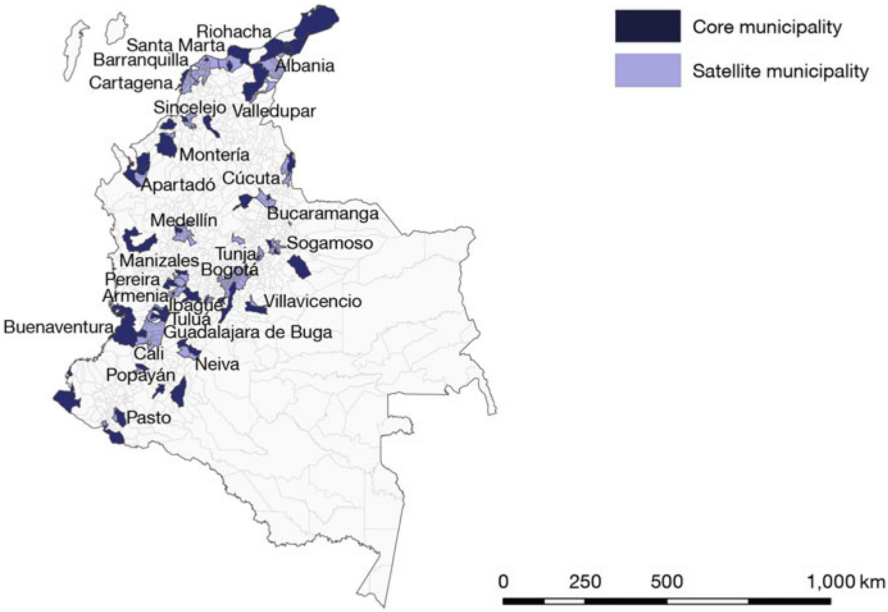


Fig. 6.2 Four regions of Colombia. (Sources: DANE and author's computations. Notes: Core municipalities in *dark blue (black)*; Satellite municipalities in *light blue (grey)*. *Narrow boundaries* between municipalities; *thick boundaries* between metropolitan areas. Metropolitan cores referenced with *large fonts*; metropolitan satellite with population above 50,000 referenced with *small fonts*)

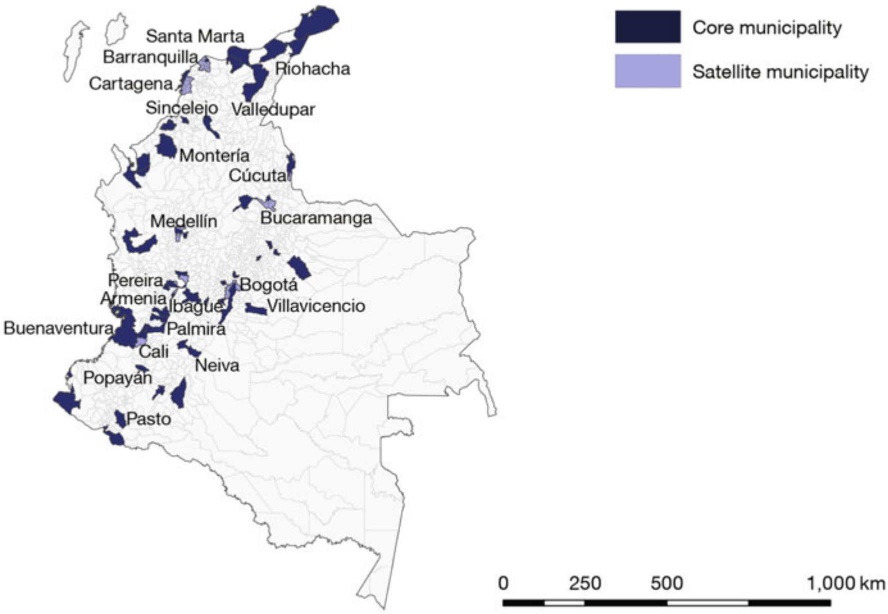
Table 6.2 Colombian urban regions with at least one metropolitan areas with population above 100,000. (Sources: DANE and author's computations.)

Urban Region	Population	Municipalities
Bogotá Fusagasugá	8,909,613	35
Carribean coast	4,139,950	37
Medellín	3,914,585	18
Cali Buenaventura Tuluá Guadalajara de Buga	3,898,886	21
Pereira Manizales Armenia Cartago	1,848,224	14
Bucaramanga Barrancabermeja	1,309,812	7
Valledupar La Guajira	1,099,054	13
Cúcuta	802,242	8
Ibagué	526,547	1
Montería	497,157	2
Villavicencio	441,906	2
Pasto	426,475	3
Apartadó Turbo	409,182	4
Sincelejo	386,560	5
Neiva	378,076	3
Sogamoso Duitama	300,580	14
Popayán	297,520	2
Tunja	206,336	6
San Andres de Tumaco	179,005	1
Florencia	157,450	1
Girardot	152,714	5
Ipiales	148,746	3
Yopal	123,361	1
Magangué	122,913	1
Quibdó	114,548	1
Lorica	114,145	1
Pitalito	113,980	1

The metropolitan areas of the Carribean coast urban region are Barranquilla, Cartagena, Santa Marta, and Zona Bananera. Despite its contiguity with Valledupar and the Guajira cities further to the north, the Sierra Nevada that separates these two areas is significant enough that they should be treated separately. The Valledupar-La Guajira region is composed of the metropolitan areas of Albania, Valledupar, Riohacha, Macaio, and Uribia. Despite its contiguity with the coffee cities of Pereira, Manizales, Armenia, and Cartago, the metropolitan area of Ibagué is left isolated because the geographical barriers that separates them is extremely important. Although the straight distance between Ibagué and Armenia is only about 50 km, the crossing of the La Línea turnpike at 3300 m often takes several hours. The completion of the tunnel between Calarcá and Cajamarca will be an important first step towards integrating Ibagué with the coffee region



Panel (a): 5% commuting threshold



Panel (b): 20% commuting threshold

Fig. 6.3 Maps of Colombian metropolitan areas with population above 100,000. (Sources: DANE and author's computations)

Medellín, and the main cities of the Coffee region: Perreira, Manizales, and Armenia. A smaller urban region also occurs around Bucaramanga and Barrancabermeja. The urban region of Bogotá contains 12 more municipalities than the previously defined metropolitan area of Bogotá but its population of 8.9 million is only marginally larger than that of metropolitan Bogotá at 8.7 million.

The second important fact that comes out of Table 6.2 is that, altogether, about 21 million people live in the four largest urban regions. This is just below half the population of the country.

We also note some interesting microfeatures about Colombian urban regions. Some regions like those around Bogotá or Medellín form highly compact regions. The urban regions of the main cities of the coffee region and that around Cali are less well formed and exhibit some 'holes'. These holes are even more apparent in the urban region of the Caribbean coast. We could choose to aggregate these unattached municipalities to the urban region that surrounds them. That would hide some interesting evolutions. These holes reveal that these urban regions are still undergoing a process of formation. The regions around Bogotá or Medellín can be thought of as already mature urban regions organised around one dominant pole. The region around Cali is still under consolidation. The same happens to the Coffee or Caribbean urban regions which have the added complication of containing several cores of relatively even populations. We can also detect potential urban regions still under formation. For instance in the Boyacá region, Duitama and Sogamoso are already integrated. Tunja, the largest metropolitan area of the region remains isolated. These two areas will eventually be integrated, perhaps even into a much larger region with Bogotá. We can also see the basis of a future urban region around Montería on the Southern part of the Caribbean region going from Magangué to Turbo.

6.5 Robustness

To show the robustness of our approach, we duplicate our main analysis for a broad range of thresholds: 1, 2, 5, 15, 20, 25, and 30 %. The two panels of Fig. 6.3 replicate the map of Fig. 6.1 for commuting thresholds of 5 and 20 %. For most large Colombian cities, a higher threshold of 20 % only makes minor differences. Of the largest 20 metropolitan areas with our preferred threshold of 10 %, 19 are still in the top 20 with a commuting threshold of 20 % and the ordering of the top 10 is unchanged. Although the metropolitan area of Bogotá loses 15 municipalities in 23 with a higher threshold of 20 %, the population remain very similar: 8.16 million instead of 8.72 million. The differences between these two rankings for the other core municipalities are even less important.

Moving to a lower threshold of 5 % also makes little difference. The ordering of the largest nine cities is unchanged. The two most important changes are the disappearance of Rionegro and Palmira which ranked 19 and 20 with a threshold of 10 %. Rionegro gets aggregated to its neighbour Medellín. The same happens to

Table 6.3 Pairwise correlations for the log population of Colombian metropolitan areas

Threshold	1 %	2 %	5 %	10 %	15 %	20 %	25 %	30 %
1 %	1							
2 %	0.991	1						
5 %	0.983	0.979	1					
10 %	0.980	0.982	0.989	1				
15 %	0.979	0.979	0.988	0.994	1			
20 %	0.979	0.978	0.987	0.994	0.999	1		
25 %	0.979	0.978	0.987	0.994	0.999	1	1	
30 %	0.979	0.978	0.987	0.993	0.999	0.999	0.999	1

Metropolitan areas with population above 50,000

Palmira with its own neighbour Cali. Interestingly, there are no other changes among the largest metropolitan areas: the three main cities of the coffee region, Armenia, Manizales, and Perreira remain separate metropolitan areas despite their proximity. Similarly the three main cities of the Caribbean coast, Barranquilla, Cartagena, and Santa Marta also remain separate.[6] These features persist even when we take an extremely low threshold of 1 %.

More generally, Table 6.3 reports log population size correlations for Colombian metropolitan areas defined according to the entire range of thresholds mentioned above. Among metropolitan areas that can be compared across thresholds (since for instance Rionegro disappears when lowering the threshold from 10 to 5 %), the correlations reported in Table 6.3 are extremely high, 0.97 or more. The correlations with our 10 % reference threshold is at least 0.98. Repeating this table using population in level or population ranks instead of log yields even higher correlations.

Next, we assess how sensitive the number of municipalities in metropolitan areas is with respect to the chosen commuting thresholds. Obviously the number of satellite municipalities is sensitive to the chosen threshold of commuting. Recall that with our reference threshold of 10 %, 99 municipalities are satellites of an urban core. With higher thresholds of 30 and 20 %, this number falls to 25 and 41, respectively. With lower thresholds of 5 and 1 %, the number of satellite municipalities increases to 180 and 616, respectively. With a threshold of 30 %, the metropolitan area of Bogotá

[6] We also start seeing satellite municipalities which are not geographically adjacent to the rest of their metropolitan areas. There are two such cases. The first is the Satanderian municipality of Sucre which get attached to Bucaramanga which is more than 200 km far. Given that this municipality is not negligibly small and sends about 7 % of its commuters to Bucaramanga, this corresponds to real flows, perhaps mostly students which are counted together with workers. The other case is Guacamayas, a tiny municipality at the North of the Boyacá region which gets attached to Bogotá which is nearly 400 km away. Given that this case is driven by only 17 ‘commuters’, this may be a statistical glitch.

Table 6.4 Spearman rank correlations for the number of satellite municipalities of metropolitan areas

Threshold	1 %	2 %	5 %	10 %	15 %	20 %	25 %	30 %
1 %	1							
2 %	0.919	1						
5 %	0.849	0.929	1					
10 %	0.747	0.802	0.865	1				
15 %	0.696	0.750	0.823	0.932	1			
20 %	0.688	0.738	0.796	0.896	0.958	1		
25 %	0.631	0.685	0.739	0.832	0.899	0.952	1	
30 %	0.598	0.641	0.695	0.781	0.847	0.904	0.948	1

Metropolitan areas with population above 50,000

has only three municipalities instead of 208 with a low threshold of 1 % even though population increases only by 37 %.[7]

To implicitly control for the large changes in the total number of satellite municipalities, we consider the Spearman rank correlation in the number of satellite municipalities as the commuting threshold varies in Table 6.4. Except for the highest thresholds for which very few metropolitan areas have satellites (only nine with a threshold of 30 %), the correlations are generally high. For instance the Spearman rank correlation between our preferred 10 % threshold and the two alternative thresholds of 5 and 20 % are 0.86 and 0.90, respectively.

Another way to assess the robustness of our findings is to look at them through the perspective of Zipf's law. This allows us to highlight the effect of the commuting threshold on the number of metropolitan areas. This exploration is also of independent interest because Zipf's law is the subject of intense academic attention. See for instance Duranton and Puga (2014) for a recent review and Pérez (2008) for a recent contribution about Colombian cities.

Since Auerbach (1913), the distribution of city sizes has often been approximated with a Pareto distribution. To do this, a popular way is to rank cities in a country from the largest to the smallest and regress log rank on log city population. Gabaix

[7] While in general municipalities that get aggregated to a core for a given threshold are also aggregated to this core or to a larger one for a lower threshold, this need not always be the case. Although exceptional, the municipality of Sutatausa provides an interesting illustration which shows the potential pitfalls of iterative aggregation. This small municipality located to the north of Bogotá sends 6 % of its workforce to San Diego de Ubaté to its north, 5 % to Tausa, 4 % to Nemocón, and 1 % to Bogotá to the south. At a 10 % threshold, Sutatausa gets aggregated to Bogotá after Tausa and Nemocón get aggregated to Bogotá. However, with a 5 % threshold, Sutatausa gets immediately aggregated to San Diego de Ubaté. Since the latter is much larger and barely sends any worker to its south, it remains an independent core with Sutatausa as satellite. This municipality of 5000 inhabitants is the only case of a satellite of Bogotá at a 10 % threshold which disappears with a 5 % threshold.

and Ibragimov (2011) highlight a possible small sample bias in the estimation of the coefficient on log city population and suggest instead using the log of the rank minus one half as the dependent variable:

$$\log(\text{Rank} - 1/2) = \beta_0 - \xi \log \text{Population} + \epsilon.$$

The estimated coefficient ξ is the shape parameter of the Pareto distribution. Zipf's law (after Zipf 1949) corresponds to the statement that $\xi = 1$. This implies that the expected size of the second largest city is half the size of that of the largest, that of the third largest is a third of that of the largest, etc.

Figure 6.4 provides a plot of the underlying data for Colombian municipalities, metropolitan areas defined according to our preferred commuting threshold of 10 %, and metropolitan areas defined with a lower threshold of 2 %.

For all Colombian municipalities in 2010, the estimated value of ξ is 0.85 suggesting a distribution more uneven than Zipf's law. We note however that this coefficient of 0.85 is mostly driven by a thin lower tail of small municipalities. It is reasonable to ignore extremely small municipalities since they are overwhelmingly rural. They are also exceptional since Colombian municipalities were designed to avoid extremely low population levels. Considering only municipalities with a population above 5000 (or 84 % of all municipalities hosting 98.7 % of the population) yields a value of ξ of 1.02 and a higher R^2 of 98 % instead of 92 % for all municipalities. To make consistent comparisons with metropolitan areas, we can restrict our attention further to only large municipalities with a population above 50,000. In this case, the estimated value of ξ is 1.07 with a R^2 of 0.99. This value of 1.07 implies less disparities in population than implied by Zipf's law. However a relatively large standard error of 0.14 makes it impossible to reject a unit coefficient and Zipf's law.[8]

For Colombian metropolitan areas defined with our preferred commuting threshold of 10 % and a minimum population size of 50,000, our estimate for ξ is 0.91 which suggests a distribution more uneven than implied Zipf's law. More generally, the estimate for ξ gets lower as lower commuting thresholds are considered. For a threshold of 30 %, we estimate $\hat{\xi}_{30} = 1.00$; for 20 % we get $\hat{\xi}_{20} = 0.95$; for 5 %, we obtain $\hat{\xi}_5 = 0.88$; for 2 % we have $\hat{\xi}_2 = 0.81$; finally for 1 %, $\hat{\xi}_1 = 0.76$. Visual inspection of Fig. 6.4 confirms this.

The counterclockwise rotation of the Zipf line as lower thresholds are considered in Fig. 6.4 is easy to understand. One the one hand, a lower commuting threshold makes the largest metropolitan areas larger. On the other hand, there are more satellite municipalities so that the number of metropolitan areas decreases. In turn, this means that the smallest areas, just above the population threshold of 50,000, have a lower rank. Hence there is a downward shift of the left tail of the Zipf's regression line when

[8] First, because the dependent variable is computed directly from the explanatory variable, measurement error on the 'true' population also affects the rank and thus leads to a downward bias for the standard errors with OLS. Gabaix and Ibragimov (2011) show that the standard error on ξ is asymptotically $\sqrt{2/n}\ \xi$ where n is the number of observations. With our data, this implies a standard error of 0.14. The values of the standard errors for the other estimates of ξ reported here are of the same magnitude.

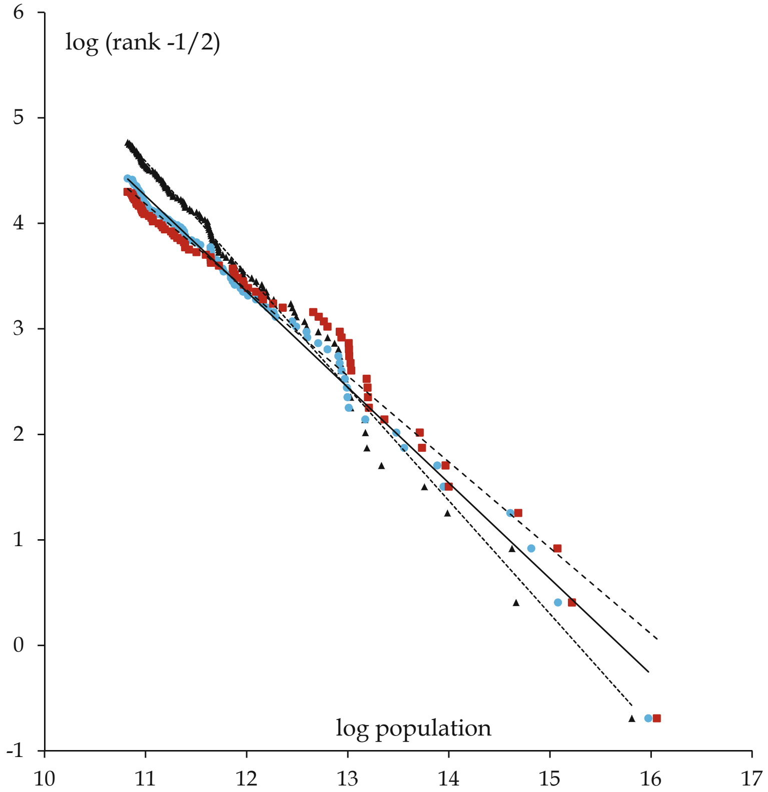


Fig. 6.4 Zipf's law for Colombian metropolitan areas and municipalities. (Sources: Author's computations with a minimum population threshold of 50,000. Notes: The *black triangles* represent municipalities and the *dotted line* is the associated regression line (slope −1.07). The *blue (light grey) dots* represent metropolitan areas defined using our preferred 10 % commuting threshold and the *plain line* is the associated regression line (slope −0.91). The *red (dark grey) squares* represent metropolitan areas defined using a 2 % commuting threshold and the *dashed line* is the associated regression line (slope −0.81))

lower commuting thresholds are used to define metropolitan areas. A combination of a shift rightwards for the largest areas and a shift downwards for the smallest areas obviously implies a flatter curve and a lower regression coefficient. We note that this would be observed even without censoring our observations at a population threshold of 50,000 since municipal aggregation overwhelmingly benefits large core municipalities and reduces the number of municipalities of a lower size.

This decline of ξ from 1.07 to 0.75 as lower commuting thresholds are considered shows that the estimates of the Pareto shape parameters for city populations are sensitive to how metropolitan areas are defined. Zipf's law is obtained exactly for a threshold of 30 % but this is arguably too high a threshold to define meaningful metropolitan areas in Colombia. This result is in contrast with older findings by

Rosen and Resnick (1980) that the size distribution of cities conforms better with Zipf's law when economically more meaningful definitions of cities are taken. This is also in contrast with more recent results by Rozenfeld et al. (2011) for the US and UK who find robust evidence for Zipf's law after defining cities using an aggregation criterion based on the geographical continuity of development.

To summarise, our findings suggest that the population of Colombian metropolitan areas is fairly insensitive to the chosen commuting threshold. As lower commuting thresholds are considered, all the metropolitan areas that remain gain population but these increases tend to be small. Relative populations are even more stable since lower thresholds lead to population gains for all metropolitan areas. By contrast, the number of satellite municipalities is more sensitive to the chosen commuting threshold. As lower thresholds are considered, the number of satellite municipalities increases dramatically. Although lower thresholds lead to more satellite municipalities for most metropolitan areas, there is also growing heterogeneity with some metropolitan areas gaining a large number of satellites and some very few. In turn, these findings suggest that the physical extent of metropolitan areas is sensitive to the chosen threshold of commuting. In turn, the aggregation of municipalities also affects estimates of the size distribution of cities. Finally, we note that the stability of both population and the number of satellite municipalities is more marked around our reference commuting threshold of 10 %.

6.6 Conclusions

In this chapter, we have proposed a simple way to define metropolitan areas relying exclusively on commuting patterns and implemented it on Colombian data. Aside from its simplicity, our approach offers two further advantages. It is fully transparent, which matters as soon as metropolitan area definitions affect policy interventions. The population of metropolitan areas is also highly robust to the details of the chosen threshold.

Our results also hold some interesting lessons for the description of large geographical networks such as labour market networks. Although their population is fairly insensitive to the details of the aggregation procedure, their physical extent is clearly much more sensitive to this. Our work also cautions against the use of simple summary statistics such as a Zipf exponent to characterise these networks.

References

Alonso, W. (1964). *Location and land use; toward a general theory of land rent*. Cambridge: Harvard University Press.

Auerbach, F. (1913). Das Gesetz der Bevölkerungskonzentration. *Petermanns Geographische Mitteilungen, 59,* 73–76.

Berry, B. J. L. (1960). The impact of expanding metropolitan communities upon the central place hierarchy. *Annals of the Association of American Geographers, 50*(2), 112–116.

Berry, B., Lobley, J., Goheen, P. G., & Goldstein, H. (1969). *Metropolitan area definition: A re-evaluation of concept and statistical practice*. Washington, DC: US Bureau of the Census.
Bode, E. (2008). Delineating metropolitan areas using land prices. *Journal of Regional Science, 48*(1), 131–163.
Briant, A., Combes, P.-P., & Lafourcade, M. (2010). Does the size and shape of geographical units jeopardize economic geography estimations? *Journal of Urban Economics, 67*(3), 287–302.
Charlot, S., & Duranton, G. (2004). Communication externalities in cities. *Journal of Urban Economics, 56*(3), 581–613.
Cheshire, P. C., & Hay, D. (1989). *Urban problems in Western Europe: An economic analysis*. London: Unwin Hyman.
Cörvers, F., Hensen, M., & Bongaerts, D. (2009). Delimitation and coherence of functional and administrative regions. *Regional Studies, 43*(1), 19–31.
Cressie, N. A. C. (1993). *Statistics for spatial data*. New York: John Wiley.
Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *Review of Economic Studies, 72*(4), 1077–1106.
Duranton, G., & Puga, D. (2014). The growth of cities. In P. Aghion & S. Durlauf (Eds.), *Handbook of economic growth* (Vol. 2, pp. 781–853). Amsterdam: North-Holland.
Fox, K. A., & Kumar, T. K. (1965). The functional economic area: Delineation and implications for economic analysis and policy. *Papers of the Regional Science Association, 15*(1), 57–85.
Gabaix, X., & Ibragimov, R. (2011). Rank-1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business Economics and Statistics, 29*(1), 24–39.
Glaeser, E. L., & Kahn, M. (2001). Decentralized employment and the transformation of the American city. *Brookings-Wharton Papers on Urban Affairs, 2*(1), 1–47.
Hall, P. G., & Hay, D. (1980). *Growth centres in the European urban system*. London: Heinemann Educational Books.
Handbury, J. H., & Weinstein, D. E. (2010). Is new economic geography right? Evidence from price data. Processed, Columbia University.
Holmes, T. J. (1999). Localisation of industry and vertical disintegration. *Review of Economics and Statistics, 81*(2):314–325.
Kanemoto, Y., & Kurima, R. (2005). Urban employment areas: Defining Japanese metropolitan areas and constructing the statistical database for them. In A. Okabe (Ed.), *GIS-based studies in the humanities and social sciences* (pp. 85–97). Boca Raton: Taylor & Francis.
Krugman, P. R. (1991). *Geography and trade*. Cambridge: MIT Press.
Marshall, A. (1890). *Principles of economics*. London: Macmillan.
Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review (Papers and Proceedings), 57*(2), 197–210.
Molina, H. (2001). *Análisis del Sistema Nacional de Ciudades. Aportes para una nueva regionalización del territorio colombiano*. New York: UNDP and Ministerio de Desarrollo Económico.
Muth, R. F. (1969). *Cities and housing*. Chicago: University of Chicago Press.
Pérez, G. J. (2008). Población y Ley de Zipf en Colombia y la Costa Caribe, 1912–1993. Documentos de trabajo sobre economía regional 71, Banco de la República.
Rosen, K., & Resnick, M. (1980). The size distribution of cities: An examination of the pareto law and primacy. *Journal of Urban Economics, 8*(2):165–186.
Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review, 101*(5):2205–2225.
Senado de la República. (2012). Informe de ponencia para secondo debate al proyecto de Ley número 141 de 2011. *Gaceta del Congreso, 137*(21), 1–16.
US Office of Management and Budget. (2010). *2010 standards for delineating metropolitan and micropolitan statistical areas; notice*. Washington, DC: Federal Register.
Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge: Addison Wesley.

Gilles Duranton is professor of real estate and holds the Dean's Chair in Real Estate at the Wharton School of the University of Pennsylvania. He obtained his PhD in economics jointly from the London School of Economics and the Ecole des Hautes Etudes en Science Sociales in Paris. He serves as co-editor for the Journal of Urban Economics and sits on the editorial board of several other academic journals. He is a fellow of the Centre for Economic Policy Research and worked as consultant on regional and urban policy for various national governments and international organisations. He was most recently the foreign advisor of a 2-year mission on urban issues for the government of Colombia. He was also the 2011 president of the North American Regional Science Association and is now vice president of the Urban Economics Association.

Chapter 7
Determinants of Business and Financial Network Formation by Japanese Start-up Firms: Does Founders' Human Capital Matter?

Hiroyuki Okamuro and Kenta Ikeuchi

Abstract Business start-ups are considered to make major contributions to economic growth. However, most lack the internal business resources necessary for survival and growth. Therefore, business and financial networks that provide business opportunities and external resources are essential for the post-entry performance of start-ups. Although preceding studies examine such networks, most do not explicitly investigate the determinants of network formation. Against this background, the present chapter argues that the formation of business and financial networks by start-up firms depends on founders' human capital, measured in terms of founders' educational attainment and business experience. This hypothesis is empirically tested using a large unique company database in Japan. Moreover, the focus is not only on the size of such networks, but also on their quality, which is measured based on the nature of major partners. The empirical results show that lengthy industry experience of 10 years or more on the part of the founder has a significant positive effect on the size of both business and financial networks, while having a university education positively affects both the size and quality of business and financial networks. The analysis further shows that a founder's specific strengths and personality traits also

This study was conducted under the project "Designing Industrial and Financial Networks to Achieve Sustainable Economic Growth" under the Ministry of Education, Culture, Sports, Science and Technology's program "Promoting Social Science Research Aimed at Solutions of Near-Future Problems." The authors are grateful to Teikoku Databank Ltd. for providing data for this study. An early version of this chapter was presented at the ZEW Conference on the Dynamics of Entrepreneurship (CoDE) in Mannheim, Germany, in October 2012 and at the RENT (Research in Entrepreneurship and Small Business) XXVI. Conference in Lyon, France, in November 2012. The authors thank the participants of these conferences for valuable comments. All remaining errors are of the authors' own responsibility.

H. Okamuro (✉)
Graduate School of Economics, Hitotsubashi University,
2-1 Naka, Kunitachi, Tokyo, Japan
e-mail: okamuro@econ.hit-u.ac.jp

K. Ikeuchi
National Institute of Science and Technology Policy (NISTEP),
3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo, Japan
e-mail: ikeuchi@nistep.go.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_7

significantly affect network formation. No distinct differences are found between the determinants of business and financial networks.

Keywords Business networks · Financial networks · Start-ups · Founders · Human capital

7.1 Introduction

Business start-ups are considered to increase competition and to contribute to innovation and economic growth (Audretsch et al. 2006). However, most start-ups are vulnerable to failure because of a lack of internal business resources. Thus, only a small proportion of start-up firms actually contribute to innovation and economic growth (Story 1994).

A firm is a pool of various tangible and intangible business resources. Both the quantity and quality of these resources determine a firm's capabilities and strongly affect its post-entry performance, including its survival and growth (Penrose 1959). Elfring and Hulsink (2003) argue that entrepreneurial networks support high-tech start-ups in their early stages to discover opportunities, secure resources, and obtain legitimacy, which is important for their survival and performance. Similarly, it can be assumed that transaction networks with business partners and financial institutions not only provide start-up firms with access to external business resources but also signal their trustworthiness to third parties. Therefore, for small start-up firms with limited internal resources, business and financial networks are essential for their survival and growth. Several studies have shown that entrepreneurial networks contribute to start-up performance. For example, Brüderl and Preisendörfer (1998) have found positive network effects on the survival and growth of new firms in Germany.

Founders of new firms often have to build and develop such networks, which is no easy task. In fact, a recent survey by the Small and Medium Enterprise Agency in Japan revealed that for the founders of start-up firms, making use of business partners and raising funds are among the most serious difficulties encountered during a firm's early stages (Small and Medium Enterprise Agency 2002). Specifically, 34 % of respondents to the survey reported that they had difficulty in gaining customers, 49 % had problems with insufficient internal funds, and 33 % had difficulty obtaining start-up finance.

Despite the importance of this issue, few empirical studies have examined the formation of business and financial networks by start-up firms. Previous studies on start-up firms regard their networks with external organizations as given. Against this background, the present chapter argues that founders' human capital plays an important role in establishing business and financial networks of start-up firms. Thus, the objective of this chapter is to empirically explore the determinants of early network formation by start-up firms, paying special attention to founders' characteristics.

While making use of business partners and obtaining external funds are two important issues for start-up firms, they have been largely addressed in different fields of literature: marketing and finance. The assumption made in the present study is that the two issues are both affected by the personal characteristics of the founders of start-up firms, but their determinants may differ. Therefore, in the analysis here, the determinants of business networks with customer firms and financial networks with banks are estimated separately and the results then compared. Moreover, the analysis will also pay attention to the quality of networks, measured in terms of the nature of major partners.

The remainder of this chapter is organized as follows. The next section provides a brief literature review to highlight the major contributions of this paper. Section 7.3 then presents the analytical framework and hypotheses, while Section 7.4 explains the empirical strategy, including the models and variables used for the estimation, and outlines the data and sample employed. Next, Section 7.5 presents and discusses the estimation results. Section 7.6 concludes this chapter.

7.2 Literature Review

Since the seminal work by Birley (1985), which explored the roles of founders' formal networks (banks, accountants, lawyers, and public agencies) and informal networks (family, friends, and business contacts), numerous studies have been conducted on the effects of entrepreneurial networks on firm performance (e.g., Cromie and Birley 1992; Ostgaard and Birley 1996; Hansen 1995; Brüderl and Preisendörfer 1998; Chell and Baines 2000; Witt 2004). In addition, Davidsson and Honig (2003) followed the activities of nascent entrepreneurs for 18 months and found that membership in business networks, such as the Chamber of Commerce, Rotary Club, or Lions Club, had a positive effect on their business performance. Furthermore, Lechner et al. (2006) argued that the "relational mix" of different types of networks is a more important factor to explain firm development than sheer network size, and demonstrated that a "reputational" network with highly regarded firms or individuals plays an important role in a firm's early stages.

However, these studies do not explicitly investigate the determinants of network formation. Moreover, they mostly focus on the social and personal ties of founders or their relationship with providers of professional services such as lawyers, accountants, and business consultants. Those studies do not explicitly consider the relationships with business partners and banks, or distinguish them from the founders' social and personal ties.

Regarding firms' networks, several empirical studies have been conducted on the determinants of research collaborations with other firms (including business partners) or research institutes; however, with the exceptions of Colombo et al. (2006) on Italy and Okamuro et al. (2011) on Japan, few studies have focused on start-up firms and the role of a founder's human capital in establishing such networks. Moreover, it appears that no studies have explicitly addressed the determinants of business networks with customers or suppliers, even though networks with business partners may strongly affect start-up performance.

There are numerous studies on the financing of small businesses and start-ups. Several studies suggest that financial constraints significantly affect start-up decisions (Evans and Jovanovic 1989) and post-entry performance (Holtz-Eakin et al.

1994; Becchetti and Trovato 2002) because, given imperfect capital markets, access to external funding such as bank loans is limited for start-up firms. Therefore, establishing transaction relationships with banks is essential for the post-entry success of start-up firms.

A number of studies have empirically investigated the effects of a founder's human capital on the external funding (especially bank borrowing) of start-up firms, with quite different results. Bates (1990) and Storey (1994) found that founders with higher educational attainment are more likely to use bank borrowing. In contrast, Astebro and Bernhardt (2003) found that business owners with higher educational attainment and rich business experience had a *lower* propensity to rely on bank loans when they started a new business. Moreover, Cassar (2004), using a sample of new businesses from a national survey in Australia, found that owner's characteristics, including sex, education, and industry experience, have *no* effects on external financing or bank loans.

However, it appears that no studies have explicitly investigated the effects of a founder's human capital on the structure of financial networks taking into account the number and nature of banks with which they have financial transactions. The number of banks with which start-up firms transact matters because it relates to the amount of bank loans available as well as to the bargaining power of start-up firms in financial transactions. Moreover, following the argument in Lechner et al. (2006), it may also be assumed that the reputation or nature of the banks is important.

Therefore, this paper aims to fill this gap by empirically examining the effects of a founder's human capital on the formation of both business and financial networks by start-up firms. Moreover, the analysis will consider not only network size, but also network quality or the reputation of network members, which is measured in term of the nature of a firm's major business partners and banks.

7.3 Analytical Framework and Hypotheses

The basic idea underlying the analysis is that a founder's human capital, including his or her education and business experience, is often the most important resource for small start-up firms. Following Fontana et al. (2006) and Okamuro et al. (2011), it is argued here that higher educational attainment (e.g., a university degree) and greater business experience represent founders' ability and personal network (in size and quality), and signal their ability to third parties such as potential business partners or lending banks.[1] It is therefore assumed that founders with higher educational attainment and longer business experience have larger business and financial networks.

[1] Along a similar vein, Bates (1990) and Story (1994) suggest that founders' education and experience may signal better human capital and hence greater viability of their start-up, enabling them to get better access to bank loans.

It is also assumed that larger networks are better for start-up firms because a larger network brings more external resources and thus more opportunities for firms. Furthermore, start-ups in larger networks are less dependent on a specific partner; reliance on a single partner makes a firm more vulnerable to exploitation. Moreover, because start-ups have limited business resources and experience, they tend to be vulnerable to external shocks, so that start-ups that have a larger number of customers are likely to be more resilient to external shocks. Therefore, one could argue that it is important for most start-up firms to increase their number of business partners.

Previous studies have focused on either business or financial networks. The assumption here is that both types of networks are affected by founders' human capital, but that the determinants may differ between them. Therefore, in the analysis below, the determinants of business and financial networks are examined separately using the same sample.

Based on these arguments, the following hypotheses regarding network size are posited:

H1a: Founders with university education tend to establish larger business networks than those without university education.

H1b: Founders with university education tend to establish larger financial networks than those without university education.

H2a: Founders with longer business experience tend to establish larger business networks than those with shorter business experience.

H2b: Founders with longer business experience tend to establish larger financial networks than those with shorter business experience.

However, it may not be just network size that is important for start-up firms. Both business partners (customer firms) and financial institutions are heterogeneous. For example, it may be more prestigious and thus more difficult for start-up firms to become suppliers for large and established firms than small, unknown customers. Similarly, it may be more prestigious and thus more difficult for start-up firms to be financed by the largest banks ("city banks") than by smaller, regional banks or credit associations,[2] because traditionally, city banks have concentrated on lending to large and established corporations. Indeed, as will be shown later, only 9 % of firms in the sample used here borrowed from a city bank. Therefore, transactions with large and established customer firms and banks, especially city banks, may indicate high-quality networks for start-up firms. Lechner et al. (2006) suggested that reputational networks with highly regarded firms play an important role for firms' development, especially at the start-up stage, by signaling their credibility.

In the following, large and established customers are regarded as superior customers and city banks as superior banks. In the same vein, business networks with superior customers are regarded as superior business networks and financial networks

---

[2] City banks are defined here as Japan's largest private banks, with branches nationwide. Although there is no legal definition of city banks, since 2006 the following four banking groups have been regarded as city banks: the Bank of Tokyo-Mitsubishi UFJ (BTMU), Sumitomo Mitsui Banking Corporation (SMBC), Mizuho Bank (MHBK), and Resona Bank. The first three groups are also called mega banks.

with superior banks as superior financial networks, and these are distinguished from other business or financial networks. Based on the above argument, superior networks are assumed to be more beneficial for start-up firms. Hence, the following hypotheses regarding the formation of superior networks are added:

H3a: Founders with university education tend to establish superior business networks to those without university education.
H3b: Founders with university education tend to establish superior financial networks to those without university education.
H4a: Founders with longer business experience tend to establish superior business networks to those with shorter business experience.
H4b: Founders with longer business experience tend to establish superior financial networks to those with shorter business experience.

It is further assumed that founders' expertise in sales or technology, as well as their personality traits also affect network formation, and thus these characteristics are controlled for in the estimation. For example, founders who have a distinct strength in sales may have an easier time finding customers, while those with accounting expertise may make a better impression on bankers.

Network formation with business partners and banks may also depend on firm-level characteristics, such as firm age, size, industry, and location. For example, it may be easier for larger firms with a longer track record to find superior customers and to obtain finance from superior banks. Therefore, these variables are also controlled for in the estimation.

7.4 Empirical Approach

This section explains the data source, the sample of firms, and the empirical models employed to estimate the determinants of the size and quality of the business and financial networks of new start-up firms.

The firm sample was obtained from the company database gathered by Teikoku Databank (TDB), the largest and oldest credit research company in Japan. The original database covers more than one million business corporations in Japan, most of which are small, unlisted firms. This database contains accounting data as well as detailed information on business transactions with other firms (business networks), financial relationships with banking institutions (financial networks), and the president of the firms. Thus, it contains all the necessary information for the present study. A distinct feature of this database is that it does not only provide the names of customer and supplier firms, but also their individual company codes, so that it is possible to accurately identify and collect data on customer firms. The same applies to the providers of bank loans. However, the database does not provide information on informal networks such as non-pecuniary transactions, political connections, and the personal ties or social networks of firm presidents.

While data on business partners are available from 2008 onward, this study uses cross-section data for the most recent year for which information on a particular start-up firm is available, since data on start-up firms are not always updated annually.

The data on start-up firms collected by TDB range from December 2007 to March 2010. It should be noted that it was not possible to obtain full data on when firms were founded, so that the most recent available data are used for the analysis. These data constraints may result in endogeneity or simultaneity problems, so that in the analysis below firm-level variables were used to control for firm-specific effects when investigating the impact of the main variables for founders' human capital.

Using this database, independent new firms in the manufacturing, wholesale, and business service sectors were identified. Firms that are less than 3 years old and still owned by the founding president are regarded here as new firms. The reason for focusing on independent firms is that firms that were spun off or set up as subsidiaries of existing firms likely can rely on the existing networks of the parent firm, while the aim of this study is to examine the networks of genuine start-up firms. The reason for focusing on these sectors is that firms in these sectors mainly transact with business firms. Therefore, all firms in consumer services are excluded. After excluding observations with a missing value, the final sample comprises 4847 firms for the analysis of business networks and 6582 firms for the analysis of financial networks.

Turning to the variables used, two groups of dependent variables are employed: proxies for network size and proxies for network quality. As for network size, this is defined as the number of customer firms when focusing on business networks and the number of banks from which a firm has borrowed when focusing on financial networks. The quality of business and financial networks is measured in terms of the number of customer firms listed on the stock exchange and whether a firm has received a loan from a city bank. Thus, in total there are four dependent variables for the empirical estimation: two representing business networks (size and quality) and two representing financial networks (size and quality).

Table 7.1 presents descriptive statistics of the firms in the sample. Regarding business networks, the average number of customer firms is 4.7 (the maximum is 35). The average number of customer firms listed on the stock exchange is about one (with a maximum of 20). Regarding financial networks, the average number of banks from which firms had obtained loans is 0.47 (with a maximum of 8); in other words, several start-ups had not taken out a loan from a bank. Moreover, only 9 % of start-up firms had borrowed from a city bank.[3]

[3] Altogether, 20 % of the start-up firms in the sample had obtained a loan from a private bank. In other words, the large majority of firms in the sample had not obtained a loan from a private bank (although some may have obtained a loan from a public bank) and about half of the firms that had obtained a loan from a private bank had done so from a city bank.

Table 7.1 Descriptive statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
(*Dependent variables*)					
Number of customers	5228	4.706	3.522	1.0	35.0
Number of listed customers	5228	1.025	1.548	0.0	20.0
Number of lending banks	7638	0.469	0.758	0.0	8.0
Borrowing from city bank	7638	0.089	0.285	0.0	1.0
(*Firm characteristics*)					
Firm age	7638	1.607	0.838	0.0	3.0
Number of employees	7638	6.4	22.1	0.0	645
Log of number of employees	7638	1.289	1.062	0.0	6.5
Manufacturing sector (reference)	7638	0.161	0.368	0.0	1.0
Wholesale sector	7638	0.457	0.498	0.0	1.0
Business service sector	7638	0.382	0.486	0.0	1.0
Location—Tokyo or Osaka	7359	0.530	0.499	0.0	1.0
(*Founder characteristics*)					
Age at start-up	6935	46.3	11.0	19.9	92.4
Sex (female)	7616	0.065	0.247	0.0	1.0
Education: University	7638	0.396	0.489	0.0	1.0
Education: Junior or technical college	7638	0.018	0.134	0.0	1.0
Industry experience: 3–10 years	7295	0.173	0.378	0.0	1.0
Industry experience: 10 years or more	7295	0.681	0.466	0.0	1.0
Managerial experience: 3–10 years	7444	0.152	0.359	0.0	1.0
Managerial experience: 10 years or more	7444	0.157	0.364	0.0	1.0
Strength in sales	7638	0.645	0.478	0.0	1.0
Strength in technology	7638	0.292	0.455	0.0	1.0
Strength in accounting	7638	0.025	0.156	0.0	1.0
Strength in administration	7638	0.204	0.403	0.0	1.0
Personality—careful	7638	0.105	0.307	0.0	1.0
Personality—active	7638	0.258	0.438	0.0	1.0
Personality—responsible	7638	0.136	0.343	0.0	1.0
Personality—technology-oriented	7638	0.108	0.311	0.0	1.0
Personality—steady	7638	0.240	0.427	0.0	1.0
Personality—visionary	7638	0.128	0.334	0.0	1.0
Personality—serious	7638	0.466	0.499	0.0	1.0

Table 7.1 (continued)

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Personality—wide personal network	7638	0.152	0.359	0.0	1.0
Personality—someone who gets things done	7638	0.330	0.470	0.0	1.0
Personality—sociable	7638	0.189	0.392	0.0	1.0
Personality—superior planning ability	7638	0.186	0.389	0.0	1.0

To examine the role of founders' human capital as a determinant of network size and quality, two variables are used: founders' educational attainment and business experience. Specifically, as a proxy for founders' educational attainment, a dummy variable indicating whether the founder had a university degree, was used. In the sample, 40 % of founders have a university degree. Next, to represent business experience, the length of a founder's experience in the same industry (industry experience) and in the management of firms (managerial experience) was used. Because this information is available only in terms of three categories—i.e., less than 3 years, 3–10 years, and 10 years or more—two dummy variables for the latter two categories are employed. In the sample, 68 % of founders had relatively lengthy industry experience of 10 years or more, but only 16 % had managerial experience of 10 years or more. It should be noted that the analysis focuses on the president as the founder of a start-up firm because of a lack of information on co-founders.[4]

In addition, two sets of control variables are included in the estimation. The first set controls for the effects of firm characteristics on network formation and consists of firm age, size, sector, and location. This study focuses on start-up firms, which are defined here as firms that are under 3 years old. Firm size is measured in terms of the natural logarithm of the number of employees. The average number of employees of firms in the sample is 6.4. Further, firms are divided into the following three sectors (with the figures in parentheses showing the share of sample firms in each sector): wholesale (46 %), business services (38 %), and manufacturing (16 %). Dummies for the first two are included in the analysis, meaning that manufacturing serves as the reference group. Finally, regarding location, firms are divided into those located in metropolitan areas and those that are not. Specifically, a dummy variable that takes a value of 1 if the firm is located in Tokyo or Osaka and 0 otherwise is used. Just over half (53 %) of all sample firms are located in these metropolitan areas.

The second set of control variables focuses on founder characteristics other than educational attainment and business experience, consisting of the founder's age, sex, forte in business, and personality. The average age of founders when they

[4] Some firms may have been founded by more than one person and thus have co-founders. Although the role of co-founders has recently attracted growing research interest, the database employed here unfortunately does not provide any information on co-founders. It only allows one to determine whether the current president is the founder, but not who among the directors are co-founders or latecomers.

started their business was 46 years. Only 6.6 % of founders were female, and a dummy for female founders is included in the analysis to control for any possible gender differences. The variable on founders' forte in business is based on the subjective assessment of TDB investigators. For the analysis below, the following categories, represented by dummy variables, are used (with figures in parentheses representing the share of founders judged to have a particular forte): expertise in sales (65 %), technology (29 %), accounting (2 %), and administration (24 %).[5] TDB investigators also assessed founders' personality traits. For the analysis below, only those traits that applied to at least 10 % of the founders of firms in the sample were used and represented using dummy variables. The traits and their percentage shares are as follows: "serious" (47 %), "someone who gets things done" (33 %), "active" (26 %), "steady" (24 %), "sociable" (19 %), "superior planning ability" (19 %), "wide personal network" (15 %), "responsible" (14 %), "visionary" (13 %), "technology-oriented" (11 %), and "careful" (11 %).[6]

Founders' characteristics and human capital were not observed and documented at the time they established their new businesses but were identified during the TDB investigation. It should be noted that what TDB collects during its investigation is information on the characteristics and human capital of the president of the firm at the time of the investigation, who may not necessarily be the founder. However, the database makes it possible to identify whether the president is also the founder of the firm, and the analysis below focuses only on those firms where this is the case. Therefore, the variables on founders' characteristics can be regarded as exogenous because they were determined before the start-up of the firms.

The correlation matrix of the dependent and independent variables is presented in the Appendix. It shows that the correlation between a firm's number of customers (which is taken to represent the size of a firm's business network) and number of lender banks (representing a firm's financial network) is relatively small at 0.145. On the other hand, the size of a firm's business network (number of customers) and the quality of its business network (as represented by the number of listed customer firms) are highly correlated (0.696), as are the size and quality of a firm's financial network as represented by the number of lender banks and whether a firm borrows from a city bank (0.531).

To estimate the determinants of network formation, negative binomial regressions are employed, except when estimating the quality of financial networks, since the relevant dependent variables are count data.[7] Regarding the quality of financial networks, a probit model is used to estimate the probability of a start-up firm obtaining finance from a city bank, since only a small proportion of start-ups used such financing, and since whether a start-up firm received a loan from a city bank is more important than how many city banks provided a loan to that firm.

---

[5] This is a multiple-choice item, so that TDB investigators could assign multiple (or no) fortes to a founder.

[6] Again, TDB investigators could assign multiple (or no) traits to a founder.

[7] The choice of negative binomial regressions is based on the results of over-dispersion tests, which suggest that such regressions rather than Poisson regressions should be used.

7.5 Estimation Results and Discussion

Table 7.2 presents the estimation results of four models with different dependent variables. Model 1 in the table corresponds to hypotheses H1a and H2a introduced above, Model 2 to hypotheses H3a and H4a, Model 3 to hypotheses H1b and H2b, and Model 4 to hypotheses H3b and H4b. For Models 1, 2, and 3 (negative binomial models), the table shows coefficient estimates, while for Model 4 (probit model) it shows marginal effects. The discussion starts by looking at the results for founders' human capital (educational attainment and business experience) and then considers the results for some of the control variables regarding founders' other characteristics as well as firm characteristics.

Beginning with the results for Model 1, which focuses on the effects on the size of business networks, the results indicate that both types of founders' human capital, that is, having received a university education and having business experience of 10 years or more in the same industry, had a significant positive effect on the number of customer firms. However, long management experience (of 10 years or more) had a significant negative effect on business network size. These results support H1a and partially support H2a (with regard to industry experience). Moreover, some of the different types of strengths and personality traits—namely, strength in sales, technology, and accounting, as well as being active, sociable, and serious—positively affected the size of a firm's business network. On the other hand, founders' age or sex had no effect on the size of business networks.

Turning to Model 2, which focuses on the quality of the business network measured in terms of the number of listed customers, the results are very similar to those for Model 1. A notable difference is that industry experience and strength in sales and accounting had no significant effects. On the other hand, as in the case of Model 1, long managerial experience had a significant negative effect. These results support H3a, but not H4a. Given the high correlation between the dependent variables of Models 1 and 2 (0.696), it is not surprising that the results for the two models are very similar.

Models 3 and 4 show the results for the size and quality of financial networks, respectively. The results indicate that university education and lengthy industry experience had a significant positive effect, which is similar to the result for business network size (Model 1). Further, lengthy managerial experience had a significant negative effect on the number of banks with which a firm transacted (Model 3). These results support H1b and partially support H2b (with regard to industry experience). As in the case of Model 1, strength in sales and accounting as well as being active, sociable, and serious had a positive effect on financial network size.

Regarding the quality of the financial network (Model 4), having a university education had a significant positive effect on the probability of borrowing from a city bank, while industry experience had no significant effect. Similar to the results for the other estimations, managerial experience had a negative impact on the probability of borrowing from a city bank. These results support H3b but not H4b.

Table 7.2 Estimation results

Model	1	2	3	4
Dependent variable	Number of customers	Number of listed customers	Number of lending banks	Borrowing from city bank
Estimation method	Negative binomial	Negative binomial	Negative binomial	Probit
(*Firm characteristics*)				
Firm age	0.087[0.012]**	0.104[0.025]**	0.419[0.023]**	0.037[0.004]**
Log of number of employees	0.146[0.009]**	0.283[0.019]**	0.247[0.017]**	0.029[0.003]**
Wholesale sector	0.062[0.030]*	– 0.085[0.063]	0.078[0.053]	0.004[0.010]
Business service sector	0.092[0.030]**	0.181[0.063]**	– 0.238[0.056]**	–0.004[0.010]
Location—Tokyo or Osaka	0.080[0.021]**	0.373[0.044]**	– 0.163[0.038]**	0.093[0.007]**
(*Founder characteristics*)				
Age at start-up	0.001[0.001]	0.004[0.002]*	– 0.002[0.002]	–0.000[0.000]
Sex (female)	– 0.068[0.045]	0.039[0.095]	– 0.398[0.095]**	–0.001[0.014]
Education: University	0.068[0.020]**	0.297[0.041]**	0.121[0.037]**	0.028[0.007]**
Education: Junior or technical college	0.097[0.069]	0.025[0.151]	0.283[0.122]*	–0.011[0.020]
Industry experience: 3–10 years	0.044[0.037]	0.016[0.079]	0.039[0.070]	0.005[0.012]
Industry experience: 10 years or more	0.076[0.032]*	0.085[0.068]	0.150[0.059]*	0.002[0.010]
Managerial experience: 3–10 years	– 0.057[0.027]*	– 0.066[0.056]	– 0.398[0.053]**	–0.021[0.007]*
Managerial experience: 10 years or more	– 0.140[0.029]**	– 0.199[0.061]**	– 0.589[0.059]**	–0.029[0.008]**
Strength in sales	0.130[0.027]**	0.018[0.056]	0.291[0.052]**	0.023[0.008]**
Strength in technology	0.067[0.028]*	0.138[0.057]*	0.054[0.052]	–0.009[0.008]
Strength in accounting	0.177[0.063]**	0.075[0.134]	0.360[0.113]**	–0.013[0.018]

Table 7.2 (continued)

Model	1	2	3	4
Dependent variable	Number of customers	Number of listed customers	Number of lending banks	Borrowing from city bank
Estimation method	Negative binomial	Negative binomial	Negative binomial	Probit
Strength in administration	0.043[0.026]	0.045[0.055]	0.109[0.048]*	0.015[0.009]
Personality—careful	– 0.038[0.032]	−0.045[0.069]	– 0.105[0.063]	−0.013[0.010]
Personality—active	0.057[0.022]*	0.104[0.047]*	0.122[0.041]**	−0.005[0.007]
Personality—responsible	– 0.032[0.026]	– 0.079[0.055]	0.084[0.048]	0.007[0.009]
Personality—technology-oriented	– 0.027[0.032]	0.092[0.065]	0.161[0.062]*	0.015[0.012]
Personality—steady	0.014[0.022]	−0.004[0.047]	0.006[0.042]	−0.007[0.007]
Personality—visionary	0.008[0.027]	0.038[0.055]	0.063[0.050]	0.010[0.009]
Personality—serious	0.045[0.021]*	−0.121[0.043]**	0.152[0.039]**	0.013[0.007]*
Personality—wide personal network	0.054[0.025]*	0.111[0.051]*	– 0.006[0.048]	0.025[0.009]**
Personality—someone who gets things done	– 0.002[0.021]	– 0.043[0.044]	0.075[0.039]	0.004[0.007]
Personality—sociable	0.095[0.023]**	0.074[0.049]	0.139[0.043]**	0.001[0.008]
Personality—superior planning ability	0.021[0.025]	0.000[0.051]	0.143[0.046]**	0.019[0.008]*
Constant	0.792[0.072]**	– 1.318[0.152]**	– 2.066[0.135]**	
Sample size	4,847	4,847	6,582	6,582
Log likelihood	– 11,711	– 6,550	– 5,881	−1,857
R-squared	0.022	0.044	0.072	0.130

The table shows regression coefficients for Models 1–3 and marginal effects for Model 4. Standard errors are in brackets
Levels of significance: $^{**}p < 0.01$; $^{*}p < 0.05$

In sum, the estimation results support hypotheses H1a, H1b, H3a, and H3b, but H2a and H2b are only partially supported (with regard to industry experience but not managerial experience), and H4a and H4b are not supported. In other words, regarding founders' human capital, having received a university education has a significant positive effect on both the size and quality of business and financial networks, while lengthy industry experience only positively affects the size of business and financial networks. On the other hand, lengthy managerial experience actually has a significant negative effect on the size and quality of both types of networks.

A possible interpretation of the unexpected negative impact of managerial experience on network formation by start-up firms is that founders with and without experience in management use different strategies. That is, founders with managerial experience may know their customer firms and banks well and prefer concentrating their business and banking relationships on a few selected partners. However, this interpretation is not consistent with the finding that founders with managerial experience also have lower quality networks, that is, they have fewer listed customers and are less likely to borrow from city banks. Another alternative interpretation is that among founders with managerial experience there are some who founded a new firm after failing with a previous one. In this case, the dummies for managerial experience partially reflect failure in a previous venture, which may act as a negative signal to potential customers and banks. This possible interpretation highlights a more general shortcoming of the analysis here, which should be noted, namely, that it cannot distinguish whether network size and quality are the result of intentional choices by founders or reflect constraints faced by founders.

Next, looking at other founder characteristics, age generally had no effect on network formation after controlling for the length of business experience. The only exception concerns the quality of business networks: older founders tend to have more listed companies as their customers. Moreover, even after controlling for their human capital, female founders tend to have fewer banks, but otherwise no significant differences by sex in network formation can be observed.

Turning to firm characteristics, the age and size (in terms of number of employees) of start-up firms, as expected, had a significant positive effect on all the dependent variables representing business and financial networks. The estimation results further suggest that the size and quality of business and financial networks differ across sectors and depend on the location of start-up firms: firms in the business services sector and in metropolitan areas have larger and better quality business networks but transact with fewer banks than firms in the manufacturing sector and in non-metropolitan areas. The probability that a firm has obtained a loan from a city bank is also higher in metropolitan areas but does not differ across sectors.[8]

---

[8] To check the robustness of these results, the sector dummies were replaced with 2-digit industry dummies and the dummy for being located in a metropolitan area (Tokyo or Osaka) with prefecture dummies. The results obtained were similar to those in Table 7.2 with regard to the effects of the other variables on network formation.

As mentioned above, the variables for founders' characteristics can be regarded as exogenous because a founder's age at start-up, gender, education level, business experiences, and personality are given before establishing a new firm; nor are they affected by network formation after start-up. However, because of data constraints it was not possible to examine any causality between firm characteristics, such as age and size, and network formation.

7.6 Concluding Remarks

Business start-ups increase competition and contribute to innovation and economic growth. However, opportunities for survival and growth are limited for most start-ups because of constraints on internal resources. Transaction networks with business partners and financial institutions can increase their opportunities for survival and growth by providing them with access to external business resources. Moreover, reputation networks with highly regarded partners may signal to third parties that the firm is capable and trustworthy, which again contributes to improving their performance.

Indeed, several studies have demonstrated that entrepreneurial networks in the early stages contribute to start-up performance. However, to the best of the authors' knowledge, few empirical studies have been carried out on the determinants of network formation by start-up firms. Thus, the purpose of the present chapter was to empirically investigate the determinants of the formation of transactions networks by independent start-up firms in Japan, using microdata compiled by TDB. The analysis distinguished between networks with business partners (business networks) and those with financial institutions (financial networks).

The key determinants on which the analysis focused are founders' educational attainment and business experience. The estimation results showed that having a university education and lengthy experience in the same industry positively affect the size of both business and financial networks, while founders' industry experience had no significant effect on whether firm transacted with highly regarded partners (listed firms and city banks). Moreover, certain business strengths and personality traits on the part of founders had a significant impact on network formation. Finally, no major differences between the determinants of business and financial network formation were found.

At this point, certain shortcomings of the analysis due to data constraints should be pointed out. First, there is a latent problem of sampling bias: the TDB database does not include all start-up firms in Japan, since TDB usually investigates firms in response to customer requests. Assuming that the investigated firms came into consideration as potential business partners or recipients of bank loans, it is therefore possible that the estimation results here are biased and are not representative of the population as a whole. Second, due of a lack of panel data the analysis relied on a cross-section sample of start-up firms within their first 3 years of operation. In this sense, the analysis is static. It would be more interesting to analyze the dynamic development of the networks of start-up firms over time.[9] Third, founders' strengths and personality traits, which partly affect network formation, were measured based on the subjective assessment of individual investigators, which is a potential shortcoming, although it is likely difficult to obtain more objective measures on those types of characteristics. Finally, no detailed information was available on transactions with each partner, such as details on transactions with banks (the interest rate on and duration of a loan, the amount of collateral involved, etc.). However, the positive effect of having a university education may be linked to differences in financing conditions, which in turn may reflect the amount of personal assets available as collateral.

Despite these caveats, the findings obtained here do offer some managerial and policy implications. For instance, founders of start-up firms and those considering to set up a new business should be aware that the likelihood of success of establishing business and financial networks, which in turn are key to the success of the start-up, may be constrained by their educational attainment and business (industry) experience. (Potential) start-up founders also need to be aware that their specific strengths and personality traits may play an important role in establishing business and financial networks. Policy makers involved in promoting start-up firms should support business and financial matching for those with lower levels of human capital, since founders with higher educational attainment and more business experience have less difficulty in forming networks.

Finally, this study points at avenues for future research on entrepreneurship. First, empirical studies on entrepreneurial networks should regard such networks as dependent on the founder's human capital. Second, future research on entrepreneurial networks should take the founder's personality traits into consideration, since these significantly affect network formation and thus the success of start-up firms.

[9] A study along these lines is the one by Schutjens and Stam (2003), who conducted a longitudinal analysis on the evolution of networks of young firms.

Appendix

Table 7.3 Correlation matrix of the variables

	Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
[1]	Number of customers	1.000									
[2]	Number of listed customers	0.696	1.000								
[3]	Number of lending banks	0.145	0.063	1.000							
[4]	Borrowing from city bank	0.075	0.054	0.531	1.000						
[5]	Firm age	0.103	0.063	0.116	0.188	1.000					
[6]	Number of employees	0.135	0.172	− 0.033	− 0.047	− 0.003	1.000				
[7]	Manufacturing sector	− 0.032	0.002	− 0.043	0.047	− 0.089	0.226	1.000			
[8]	Wholesale sector	0.002	− 0.026	0.015	0.057	− 0.024	− 0.095	− 0.375	1.000		
[9]	Business service sector	0.024	0.024	0.020	− 0.093	0.095	− 0.090	− 0.441	− 0.666	1.000	
[10]	Location—Tokyo or Osaka	0.095	0.113	0.120	− 0.079	0.060	0.025	− 0.236	− 0.080	0.267	1.000
[11]	Age at start-up	0.006	0.064	− 0.082	− 0.019	− 0.059	0.186	0.281	0.152	− 0.373	− 0.220
[12]	Sex (female)	0.008	− 0.000	0.001	0.039	0.008	− 0.029	0.027	− 0.032	0.009	0.054
[13]	Education: University	− 0.033	0.023	0.044	− 0.085	− 0.024	0.025	0.008	0.025	− 0.030	0.079
[14]	Industry experience: 3–10 years	0.045	− 0.038	0.017	0.009	0.033	− 0.068	− 0.089	− 0.101	0.169	0.087
[15]	Industry experience: 10 years or more	0.037	0.073	− 0.018	0.016	− 0.007	0.079	0.077	0.102	− 0.161	− 0.061
[16]	Managerial experience: 3–10 years	0.108	0.163	− 0.008	− 0.013	0.038	0.166	− 0.019	0.008	0.007	0.047
[17]	Managerial experience: 10 years or more	− 0.038	− 0.024	− 0.044	− 0.013	0.000	0.026	0.168	0.050	− 0.183	− 0.137
[18]	Strength in sales	0.058	0.003	0.061	0.070	0.047	− 0.042	− 0.180	0.245	− 0.093	0.032

Table 7.3 (continued)

	Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
[19]	Strength in technology	− 0.015	0.038	0.009	− 0.053	0.018	− 0.049	0.015	− 0.242	0.222	0.109
[20]	Strength in accounting	− 0.002	− 0.020	− 0.049	0.013	− 0.012	− 0.003	0.100	− 0.028	− 0.053	0.017
[21]	Strength in administration	0.009	0.003	− 0.056	− 0.027	− 0.017	0.072	0.158	− 0.036	− 0.092	− 0.094
[22]	Personality—careful	− 0.016	0.019	− 0.055	− 0.063	0.011	0.086	0.122	− 0.030	− 0.069	− 0.099
[23]	Personality—active	0.077	0.083	0.021	0.029	0.054	0.051	− 0.058	0.056	− 0.007	0.049
[24]	Personality—responsible	− 0.060	− 0.040	0.002	0.009	0.031	0.071	0.111	0.011	− 0.100	0.027
[25]	Personality—technology-oriented	− 0.057	0.004	0.002	0.008	0.045	− 0.049	0.073	− 0.147	0.084	0.041
[26]	Personality—steady	0.008	0.009	0.047	0.082	− 0.039	0.099	0.058	0.109	− 0.152	− 0.072
[27]	Personality—visionary	− 0.001	− 0.035	0.003	− 0.008	0.019	0.004	− 0.008	− 0.081	0.084	0.049
[28]	Personality—serious	− 0.026	− 0.058	− 0.015	0.035	0.012	− 0.008	0.060	0.050	− 0.097	− 0.127
[29]	Personality—wide personal network	0.023	0.075	0.079	0.022	− 0.015	− 0.032	− 0.057	0.019	0.027	0.069
[30]	Personality—someone who gets things done	0.037	− 0.013	0.036	0.010	0.065	− 0.007	− 0.091	0.038	0.037	0.139
[31]	Personality—sociable	− 0.031	− 0.033	0.003	0.005	− 0.032	− 0.030	− 0.062	0.091	− 0.038	− 0.038
[32]	Personality—superior planning ability	0.029	− 0.018	− 0.019	− 0.072	0.027	− 0.042	− 0.083	− 0.154	0.216	0.265

Table 7.4 Correlation matrix of the variables (cont.)

	Variable	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]
[11]	Age at start-up	1.000										
[12]	Sex (female)	− 0.045	1.000									
[13]	Education: University	0.039	− 0.152	1.000								
[14]	Industry experience: 3–10 years	− 0.416	0.018	− 0.024	1.000							
[15]	Industry experience: 10 years or more	0.441	− 0.083	0.041	− 0.763	1.000						
[16]	Managerial experience: 3–10 years	0.023	− 0.004	0.028	0.006	0.030	1.000					
[17]	Managerial experience: 10 years or more	0.396	0.040	− 0.033	− 0.182	0.201	− 0.231	1.000				
[18]	Strength in sales	− 0.077	− 0.097	0.027	0.010	0.011	− 0.061	− 0.036	1.000			
[19]	Strength in technology	− 0.072	− 0.016	− 0.027	− 0.023	0.078	− 0.019	− 0.054	− 0.460	1.000		
[20]	Strength in accounting	− 0.032	0.106	− 0.043	0.013	− 0.111	0.060	− 0.044	− 0.153	− 0.078	1.000	
[21]	Strength in administration	0.138	0.059	0.057	− 0.036	− 0.016	0.108	0.142	− 0.303	− 0.222	0.074	1.000
[22]	Personality—careful	0.104	− 0.018	0.015	− 0.007	0.007	− 0.014	0.009	− 0.041	0.017	− 0.004	0.049
[23]	Personality—active	− 0.035	0.017	− 0.025	0.032	− 0.001	0.052	0.081	0.158	− 0.116	− 0.036	− 0.037
[24]	Personality—responsible	0.037	− 0.033	− 0.026	− 0.111	0.062	0.042	0.023	− 0.025	− 0.024	0.076	0.102
[25]	Personality—technology-oriented	0.008	0.028	0.015	− 0.002	0.049	− 0.077	− 0.038	− 0.289	0.437	− 0.022	− 0.142
[26]	Personality—steady	0.109	− 0.045	0.038	− 0.074	0.070	− 0.041	− 0.011	− 0.032	− 0.036	0.016	0.064
[27]	Personality—visionary	− 0.093	− 0.007	0.016	0.068	− 0.039	0.022	− 0.052	0.089	0.047	0.042	− 0.025
[28]	Personality—serious	0.049	− 0.042	0.042	− 0.020	0.005	− 0.013	− 0.044	− 0.122	0.056	0.032	0.092
[29]	Personality—wide personal network	0.028	− 0.032	− 0.016	0.019	− 0.002	0.010	− 0.002	0.119	− 0.041	− 0.049	0.005

Table 7.4 (continued)

	Variable	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]
[30]	Personality—someone who gets things done	− 0.100	− 0.016	− 0.015	− 0.027	0.018	0.049	0.081	0.132	− 0.064	− 0.013	− 0.005
[31]	Personality—sociable	− 0.003	0.032	− 0.052	0.012	0.014	− 0.024	− 0.034	0.127	− 0.076	− 0.048	− 0.026
[32]	Personality—superior planning ability	− 0.181	0.082	0.021	0.045	− 0.065	0.026	− 0.027	0.036	0.061	0.015	0.001
	Variable	*[22]*	*[23]*	*[24]*	*[25]*	*[26]*	*[27]*	*[28]*	*[29]*	*[30]*	*[31]*	*[32]*
[22]	Personality—careful	1.000										
[23]	Personality—active	− 0.131	1.000									
[24]	Personality—responsible	0.021	− 0.044	1.000								
[25]	Personality—technology-oriented	0.076	− 0.132	− 0.071	1.000							
[26]	Personality—steady	0.052	− 0.211	0.008	− 0.020	1.000						
[27]	Personality—visionary	− 0.086	− 0.016	− 0.018	0.022	− 0.107	1.000					
[28]	Personality—serious	0.141	− 0.289	0.084	0.084	0.211	− 0.171	1.000				
[29]	Personality—wide personal network	− 0.102	0.058	− 0.081	− 0.058	− 0.082	0.040	− 0.171	1.000			
[30]	Personality—someone who gets things done	− 0.129	0.115	− 0.048	− 0.071	− 0.131	0.040	− 0.188	0.051	1.000		
[31]	Personality—sociable	− 0.043	0.046	− 0.091	−0.075	− 0.044	− 0.016	− 0.105	0.003	− 0.030	1.000	
[32]	Personality—superior planning ability	− 0.123	0.010	− 0.016	− 0.046	− 0.193	0.154	− 0.225	0.027	0.181	− 0.033	1.000

References

Astebro, T., & Bernhardt, I. (2003). Start-up financing, owner characteristics, and survival. *Journal of Economics and Business, 55,* 303–319.

Audretsch, D. B., Keilbach, M. C., & Lehmann, E. E. (2006). Entrepreneurship and economic growth. New York: Oxford University Press.

Bates, T. (1990). Entrepreneur human capital inputs and small business longevity. *Review of Economics and Statistics, 72,* 551–559.

Becchetti, L., & Trovato, G. (2002). The determinants of growth for small and medium sized firms. The role of the availability of external finance. *Small Business Economics, 19,* 291–306.

Birley, S. (1985). The role of networks in the entrepreneurial process. *Journal of Business Venturing, 1,* 107–118.

Brüderl, J., & Preisendörfer, P. (1998). Network support and the success of newly founded firms. *Small Business Economics,10,* 213–225.

Cassar, G. (2004). The financing of business start-ups. *Journal of Business Venturing, 19,* 261–283.

Chell, E., & Baines, S. (2000). Networking, entrepreneurship and microbusiness behaviour. *Entrepreneurship and Regional Development, 12,* 195–215.

Colombo, M. G., Grilli, L., & Piva, E. (2006). In search of complementary assets: The determinants of alliance formation of high-tech start-ups. *Research Policy, 35,* 1166–1199.

Cromie, S., & Birley, S. (1992). Networking by female business owners in Northern Ireland. *Journal of Business Venturing, 7,* 237–251.

Davidsson, P., & Honig, B. (2003). The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing, 18,* 301–330.

Elfring, T., & Hulsink, W. (2003). Networks in entrepreneurship: The case of high-technology firms. *Small Business Economics, 21,* 409–422.

Evans, D. S., & Jovanovic, B. (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy, 97,* 808–827.

Fontana, R., Geuna, A., & Matt, M. (2006). Factors affecting university-industry R&D projects: The importance of searching, screening, and signaling. *Research Policy, 35*, 309–323.

Hansen, E. L. (1995). Entrepreneurial networks and new organization growth. *Entrepreneurship Theory and Practice, 19,* 77–79.

Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1994). Sticking it out: entrepreneurial survival and liquidity constraints. *Journal of Political Economy, 102,* 53–75.

Lechner, C., Dowling, M., & Welpe, I. (2006). Firm networks and firm development: The role of the relational mix. *Journal of Business Venturing, 21,* 514–540.

Okamuro, H., Kato, M., & Honjo, Y. (2011). Determinants of R&D cooperation in Japanese start-ups. *Research Policy, 40*, 728–738.

Ostgaard, T. A., & Birley, S. (1996). New venture growth and personal networks. *Journal of Business Research, 36,* 37–50.

Penrose, E. T. (1959). *The theory of the growth of the firm*. London: Basil Blackwell.

Schutjens, V., & Stam, E. (2003). The evolution and nature of young firm networks: A longitudinal perspective. *Small Business Economics, 21,* 115–134.

Small and Medium Enterprise Agency (2002). White paper on small and medium enterprises in Japan (in Japanese). Tokyo: Gyosei.

Storey, D. J. (1994). The role of legal status in influencing bank financing and new firm growth. *Applied Economics, 26,* 129–136.

Witt, P. (2004). Entrepreneurial networks and the success of start-ups. *Entrepreneurship and Regional Development, 16,* 391–412.

Hiroyuki Okamuro is professor at the Graduate School of Economics at Hitotsubashi University and visiting research fellow at the National Institute of Science and Technology Policy (NISTEP). His main field of research is empirical industrial organization, especially innovation and entrepreneurship. He has published papers in the *Journal of Economic Behavior and Organization*,

the *Journal of Banking and Finance, Research Policy, Economics of Innovation and New Technology*, the *Journal of Technology Transfer, Small Business Economics*, and the *Journal of Small Business Management*, among others. Since 2013 he has served as an associate editor of the Journal of Small Business Management. He is a vice president of Japan Academy of Small Business Studies (JASBS) and a vice president of Asian Council for Small Business (ACSB). He received a B.A. and M.A. in Economics from Hitotsubashi University and a Ph.D. in Economics (Dr. rer. pol.) from University of Bonn, Germany.

Kenta Ikeuchi is research fellow at the National Institute of Science and Technology Policy (NISTEP). He also served as visiting lecturer at the Faculty of Economics at Nihon University and as visiting researcher at the National Graduate Institute for Policy Studies (GRIPS). His main fields of research are business economics, spatial economics and innovation policy. He has published a paper with Hiroyuki Okamuro in the refereed book *Entrepreneurship, Growth and Economic Development* (*Frontiers in European Entrepreneurship Research 23*), edited by M. Raposo et al., *Edward Elgar Publishing Limited*, 2011, among others. He received a B.A. in Commerce from Chuo University and a M.A. in Economics from Hitotsubashi University.

Chapter 8
Geographical Spread of Interfirm Transaction Networks and the Great East Japan Earthquake

Yukiko Umeno Saito

Abstract This chapter examines how the impact of the Great East Japan Earthquake spread geographically to unaffected areas through interfirm transaction networks. Previous studies on transaction networks revealed small-world structure and geographical proximity, which have conflicting implications for the geographical impact of the earthquake. Using interfirm transaction data from approximately 800,000 firms, it is examined how firms in physically unaffected areas are linked with those in the affected areas. It is found that only 3 % of firms in unaffected areas have direct transaction links with those in the affected areas. On the other hand, the share increases to 40–60 % if indirectly linked transaction partners (i.e. partners of partners) are taken into account. Further, it is shown that it is a small number hub firms with interfirm links spanning larger distances that are responsible for linking more local networks in different regions and hence for geographically spreading the economic impact of the earthquake.

Keywords Interfirm networks · Geographical proximity · Earthquake

8.1 Introduction

Following the Great East Japan Earthquake in March 2011, many firms, including those in physically unaffected areas, re-evaluated the importance of understanding the structure of their supply chain. In normal times, not many firms are likely to formulate management strategies taking into account firms with which they are indirectly linked (partners of partners). However, the competitiveness of Japanese industry is widely believed to depend on strong linkages between transaction partners. The experience following the earthquake has shown that this competitiveness is vulnerable to disruptions of such linkage. Against this background, Tokyo Shoko Research (TSR) conducted a survey entitled "Research Related to the Great East

Y. U. Saito (✉)
Research Institute of Economy, Trade and Industry,
1-3-1 Kasumigaseki Chiyoda-ku, Tokyo, Japan
Tel.: +81-3-3501-8312
e-mail: saito-yukiko@rieti.go.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_8

Japan Earthquake." They defined the affected areas and analyzed the population and properties of firms in those areas to estimate the impact of the damage. According to their analysis, there are 32,341 firms in the physically affected areas, with a total of 363,796 employees. However, as is well known, many more firms were impacted by the earthquake via supply chains.

This chapter examines how the impact of the Great East Japan Earthquake spread geographically to unaffected areas through interfirm transaction networks. Important aspects in analyzing the extent of the impact of the earthquake through such networks are geographical proximity and the structure of interfirm transaction networks, since these two aspects greatly affect the percentage of firms indirectly affected by the disaster (those in physically unaffected areas).

It is widely recognized in the social sciences that many networks, including interfirm networks, have small-world characteristics, meaning that a large number of network members are linked with each other either directly or indirectly. In fact, a well-known phenomenon in the field of network science is that many of the networks observed in society have common structural characteristics. For instance, a phenomenon common to many networks is that they are self-organizing. Another is that many networks are both scale free and have small-world characteristics. (For details, see, e.g., Albert and Barabási 2002). If, as in other networks, small-world characteristics are observed in interfirm transaction networks, this would imply that an extremely high number of firms in unaffected areas could be damaged through transaction links without firms being aware of this risk. Ohnishi et al. (2010) calculated the shortest path length for every pair of firms in interfirm transaction networks and found that the average such length was extremely short. This means that the networks had a small-world structure. Given this finding, one would expect that many firms outside the affected areas are connected with firms in the affected areas through interfirm transaction relationships.

Further, it is well known that interfirm transactions are one of the key factors giving rise to industry agglomeration. To reduce transaction costs, firms tend to locate close to transaction partners, as shown both in theoretical studies (e.g. Duranton and Puga 2004) and empirical ones (e.g. Rosenthal and Strange 2001; Ellison et al. 2010; Nakajima et al. 2013). Examining transaction relationships in Japan, Nakajima et al. (2012) found that transactions tend to take place within relatively small areas. Specifically, they calculated the geographical distances between transaction partners for manufacturing firms and found that half of the transaction relationships were with partners within a 40 km radius. However, if firms' transaction relationships are concentrated within close geographical proximity, this would imply that the geographical spread of supply chain disruption as a result of damage from the earthquake should be limited. Thus, the small-world structure of networks on the one hand and geographical proximity in interfirm transaction links on the other appear to have conflicting implications in terms of the geographical impact of the earthquake, meaning that in order to resolve this apparent conflict, geographical property of indirect transaction links need to be taken into account.

To do so, affected areas in this chapter are defined as in the aforementioned survey by TSR. Data on interfirm transaction are then used to identify direct and indirect transaction partners of firms in affected areas. This analysis shows that less than 3 % of firms in the unaffected areas had transaction partners in affected areas. This share, however, rises to 40–60 % once indirect transaction relationships (partners of partners) are taken into account. The share rises further to 90 % once longer path lengths (partners of partners' partners) are taken into account. The latter result indicates that most firms in the unaffected areas have indirect relations with firms in the affected areas.

This chapter further examines the geographical proximity of direct and indirect transactions. It is found that the median distance between indirect partners is 255 km, which is much larger than that between direct partners (29 km). This means that indirect transactions are widely dispersed. In addition, in order to investigate the role of hub firms, a counterfactual network without hub firms is considered and it is found that a few hub firms greatly increase the geographical spread of networks.

The remainder of this chapter is organized as follows. The next section provides an introduction to network analysis, while Section 8.3 explains the data used in the analysis here. Section 8.4 then presents the results of the analysis and Section 8.5 concludes.

8.2 Network Analysis

Network analysis has its origin in natural sciences, but has subsequently come to be used also in sociology, management studies, and other areas. In graph theory, "graphs" are described by sets of "nodes" and "edges" (or "links"). In Fig. 8.1, circles represent nodes and lines symbolize edges. Each node is connected by edges, and various combinations of nodes and edges produce different characteristics.

In the natural sciences, various network structures have been analyzed, and graph theory is used to categorize networks in terms of their patterns. For example, common structures have been observed in networks such as ecosystems (e.g., protein interactions and metabolic pathways) and the Internet. These structures are often captured by the number of links and shortest path length. For example, in the Fig. 8.1, the number of links for nodes A, B, C and D are 3, 2, 5 and 1, respectively. The shortest path length between A and B is 1; that between A and C is 2, and that between A and D is 3.

One widely observed aspect regarding network structure is that many of them are scale-free. Such networks have a power-law distribution for the number of links. Power-law distributions are frequently observed for phenomena in sociology. The distribution of wealth is an often-cited example. That is, it is frequently observed that 20 % of the population hold 80 % of the wealth. The fact that the number of links in an interfirm network follows a power-law distribution means that there are hub firms with an extremely high number of links.[1] Another widely observed aspect regarding

[1] A review of power-law distributions in the fields of economics and finance is provided by Gabaix (2009).

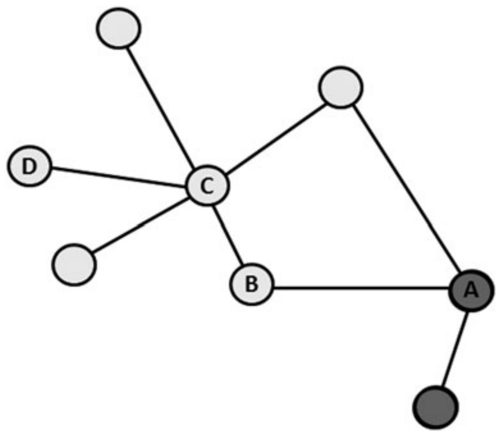


Fig. 8.1 Example of a graph

network structure is that many networks display a small-world structure. This means that the path length between any two nodes is extremely small, as illustrated, for example, by the well-known observation that there are only 6° of separation between any two people in the world.

In the field of economics and management studies, network analysis is often used as a way to visualize relationships. For example, network analysis can help to trace the transaction partners and partners' partners of specific firms to predict transmission paths of shocks, such as a downturn in corporate performance or bankruptcy. As an example, we consider how the impact of the Great East Japan Earthquake on businesses was transmitted through interfirm transaction networks. For illustration, gray circles in Fig. 8.1 represent firms in affected areas and white circles firms in unaffected ones. In the analysis below, firms in the affected areas will be referred to as Tier 0 firms, while the transaction partners of Tier 0 firms are referred to as Tier 1 firms. Hence, the partners of Tier n firms are defined as Tier n + 1 firms, if the firms are not defined as lower Tier firms. In defining these firms, there can be no overlap with firms in other tiers; thus, a Tier n firm cannot be in Tier m where m is larger than n. In Fig. 8.1, firms A, B, C and D are Tier 0, 1, 2 and 3 firms, respectively.

8.3 Data

The data used are from a database created by TSR and containing approximately 800,000 firms, including many small and medium-sized firms. This database covers about half of the registered firms in Japan, making it very comprehensive. The survey was conducted in 2006. The data include various types of information on firms, such as their industry classification, year of establishment, sales and profits for the preceding 3 years, number of employees, address, and up to 24 partners (suppliers, buyers, and major shareholders). Transaction partners are identified by their identification code, so that the information on transaction partners can be merged with the firm information. The database contains approximately 4 million transaction relationships.

Further, using the address-matching service provided by the Center for Spatial Information Science at the University of Tokyo,[2] firm addresses in the TSR database were geo-coded in terms of longitude and latitude. This makes it possible to trace the geographical distribution of firms and calculate distances between transaction partners. Given that the TSR database contains only up to 24 transaction partners for each firm, not all transaction relationships are included in the data. However, counting the number of times firms were listed as transaction partners following Saito and Watanabe (2007), it is possible to identify hub firms with more than 10,000 relationships. Employing this approach, Saito and Watanabe (2007) showed that the interfirm network is scale-free, while Ohnishi et al. (2010) showed that it is a small-world network.

The focus of the present analysis is on supply chains, so links here consist of links between customers and suppliers. The transaction relationship data contain information on the flow of goods, so the corresponding graph is a directed one (i.e., one that has link directions). It should be noted, however, that link direction is not considered in the analysis here.

8.4 Geographical Spread of the Earthquake

Let us start by defining the affected areas and showing the geographical distribution of firms in those areas. Following this, the analysis looks at the ratio of indirect transaction partners in unaffected areas and highlights some interesting feature of these ratios. Then, it is examined how direct and indirect transactions are geographically distributed. Finally, the connecting paths of indirect partners are identified and the role of regional hub firms in propagating the impact of the earthquake through supply chains is discussed.

8.4.1 Definition of Affected Areas

The definition of affected areas here follows that in the TSR survey. Specifically, the affected areas consist of 44 municipalities along the Pacific coast in Aomori, Iwate, Miyagi, and Fukushima prefectures. Using the longitude and latitude of firms' location, Fig. 8.2 maps firms in the affected areas.

---

[2] See http:/newspat.csis.u-tokyo.ac.jpgeocode/.

Fig. 8.2 Geographical distribution of tier 0 firms

8.4.2 Share of Indirect Transaction Partners

This section examines the number of indirect transaction partners and presents the percentage of firms in each tier by region. As mentioned, firms in the affected areas are referred to as Tier 0 firms, their transaction partners as Tier 1 firms, and partners' partners as Tier 2 firms. Table 8.1 shows the cumulative percentages of firms.[3] That is, the first column shows the percentage of Tier 0 firms, the second the sum of the percentage of Tier 0 and 1 firms, and so forth.

The affected areas are in the Tohoku region, so Tier 0 firms are restricted to that region. Tier 0 firms, that is, firms in the affected areas, make up 17 % of all firms in the Tohoku region. Tier 0 firms and their transaction partners, i.e., Tier 1 firms, together make up 34 % of all firms in the Tohoku region, and this share rises to 82 % when including partners' partners (i.e., Tier 2 firms). This shows that the large majority of firms in the Tohoku were indirectly affected.

Focusing on the Kanto region, which has largest number of firms of any region, only 3 % of firms transacted directly with firms in the affected areas, but counting partner's partners, the share rises to 58 %, i.e., more than half. Even in areas far from

[3] To provide more detailed information, cumulative percentages for each prefecture and industry are shown in the Appendix.

Table 8.1 Cumulative percentage of firms

	Tier 0 (%)	Tier 1 (%)	Tier 2 (%)	Tier 3 (%)	Tier 4 (%)	Tier 5 (%)
Total	2	5	57	90	96	97
Hokkaido	0	2	60	96	99	99
Tohoku	17	34	82	97	98	99
Kanto	0	3	58	89	95	95
Chubu	0	1	52	91	97	98
Kinki	0	1	54	88	95	96
Chugoku/Shikoku	0	1	47	90	97	97
Kyushu	0	0	43	88	97	97

the affected areas, e.g., the Kyushu area, nearly half of all firms were partners of partners in the affected areas.

Regarding Tier 3 firms, the percentage of firms in all regions increases to around 90 %, showing that there were few firms without direct or indirect links with firms in the affected area. Thus, the network of interfirm transactions in Japan is truly small-world. This can also be visually seen in Fig. 8.3, which shows the cumulative percentages of firms in each prefecture linked to firms in the affected areas. The values are shown in Table 8.8 in the Appendix.

8.4.3 Geographical Proximity of Indirect Transactions

The most interesting finding in the previous sub-section is that in the unaffected areas the population of Tier 2 firms is much larger than that of Tier 1 firms. To understand the underlying mechanism for this, this section examines the geographical proximity of direct and indirect links, focusing on Japan as a whole rather than firms in the areas affected by the earthquake.

Direct transaction relationships tend to be localized, as shown in Nakajima et al. (2012). Specifically, they found that half of the direct transactions of manufacturing firms are within 40 km. In this section, it is found that more than half of direct transactions of firms of all industries are within 29 km (Table 8.2). This suggests that transactions of non-manufacturing firms are more localized than those of manufacturing firms. The fact that the percentages of Tier 1 firms in the unaffected areas in Table 8.1 are extremely small is in line with the observation that direct transaction relationships tend to be localized.

However, the fact that the percentages of Tier 2 firms in the unaffected areas are extremely large suggests that indirect transactions are not localized. To examine the geographical proximity of indirect transactions, Table 8.2 shows the distances between indirect transaction partners, i.e., partners of partners. The figures indicate

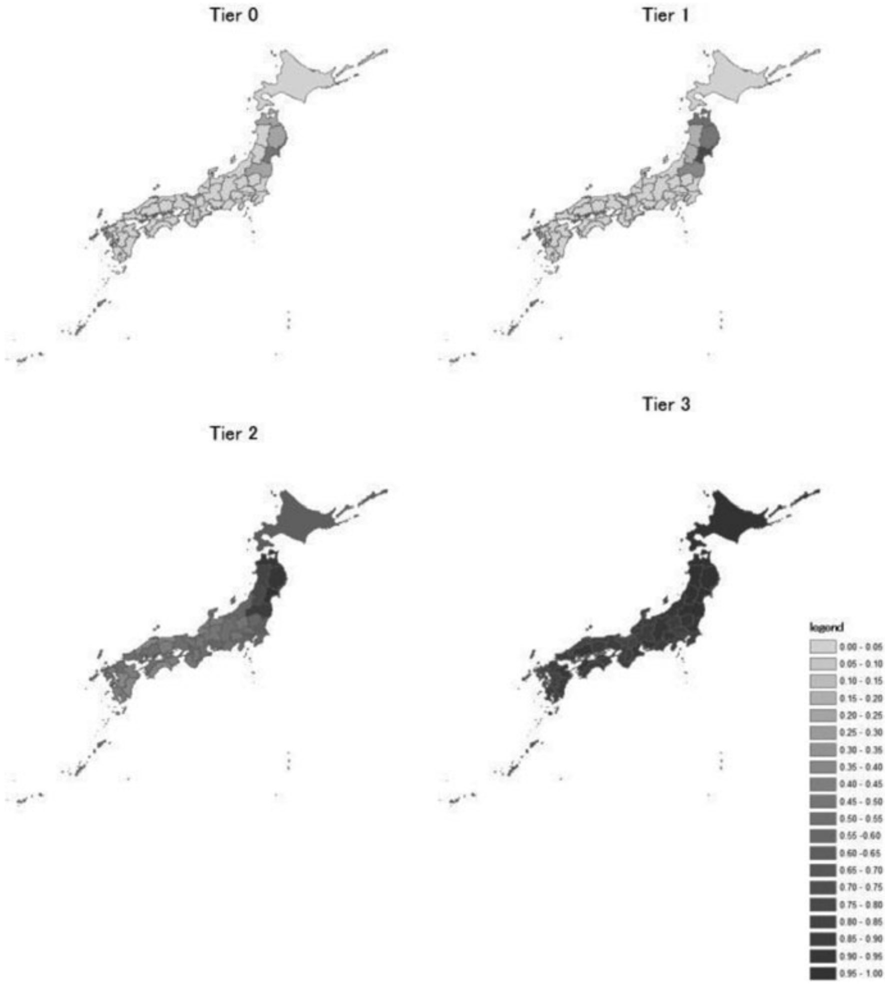


Fig. 8.3 Cumulative percentage of firms by prefecture

that the median distance between indirect partners is 255 km, which is much larger than the distance between direct partners (29 km).

Furthermore, to examine how widely firms that are linked indirectly are dispersed, the distance between firms and partners of partners and that between randomly chosen firms are compared. Retaining the number of transaction links for each firm, but randomly shuffling their indirect transaction partners, the median distance between randomly chosen partners is 416 km. Thus, the median distance between firms and partners of partners is approximately 9 times as large as that between direct transaction partners, but much similar to that between randomly chosen firms.

Table 8.2 Geographical proximity of indirect transactions

	No. of obs.	Mean	Std. dev.	p50	p25	p75
Direct	34,61,510	169.64	271.73	28.84	6.74	231.67
Indirect	17,50,0000	341.83	340.19	255.44	40.83	504.74
Random	17,50,0000	481.68	378.26	416.37	186.24	717.41

Table 8.3 Geographical links between tier 1 and 2 firms

		Tier 2						
		Hokkaido (%)	Tohoku (%)	Kanto (%)	Chubu (%)	Kinki (%)	Chugoku/ Shikoku (%)	Kyushu (%)
Tier 1	Hokkaido	75	1	5	4	3	2	1
	Tohoku	8	80	10	8	7	7	4
	Kanto	11	14	50	24	29	20	15
	Chubu	1	1	8	40	10	5	3
	Kinki	3	3	13	13	33	11	8
	Chugoku/ Shikoku	1	1	7	5	10	45	6
	Kyushu	1	0	7	5	8	9	63

8.4.4 Connecting Path of Indirect Links

Returning to the analysis of how firms are linked to firms in the areas affected by the earthquake, this section examines the path of indirect links, that is, how Tier 1 and 2 firms are geographically connected. First, Fig. 8.4 shows the geographical distribution of Tier 1 firms. The figure shows that these firms are widely distributed throughout the country, although, as shown in Table 8.1, the percentage of firms is extremely small. Next, Table 8.3 shows how Tier 1 firms are connected to Tier 2 firms by region. For each Tier 2 firm in a region, the region of Tier 1 firms that connect it to Tier 0 firms is identified, and the percentage of Tier1 firms by region calculated. Looking at Table 8.3, it can be seen that the percentages are largest along the matrix diagonal, which means that Tier 2 firms are typically connected to Tier 0 firms through Tier 1 firms in the same region. For example, 75 % of Tier 2 firms in Hokkaido are connected to Tier 0 firms through Tier 1 firms in that region. This result suggests that certain Tier 1 firms which transact over long distances play an important role in linking Tier 0 and 2 firms, explaining the high share of Tier 2 firms in Table 8.1.

Fig. 8.4 Geographical distribution of tier 1 firms

8.4.5 The Role of Regional Hub Firms in Spreading the Impact of the Earthquake

This section examines the role of Tier 1 firms in the unaffected areas in greater detail. The previous sub-section suggested that a large share of Tier 2 firms in the unaffected areas is linked with Tier 0 through Tier 1 firms in the same region. Given that the number of Tier 1 firms is much smaller than that of Tier 2 firms, it can be presumed that hub firms in each region play an important part in indirectly linking firms across larger distances.

To investigate this issue more close, his section starts by examining the relationship between firm characteristics and the distances between transaction partners. The analysis focuses on hub firms, for which three alternative definitions are used: firms with at least 100, 500, and 1000 transaction partners. From all the transaction relationships in the sample, those in which at least one of the firms was a hub firm are extracted, and the distances calculated. The results are shown in Table 8.4, with the distances for all links shown for comparison. They show that transaction distances were longer for transactions involving bigger hub firms. This finding is in line with the conjecture that Tier 2 firms are linked to Tier 0 firms in the affected areas via hub firms.

In other words, the pattern that results in many Tier 2 firms in unaffected areas being linked to Tier 0 firms appears to be as follows. Regional hub firms in the unaffected areas have links spanning a considerable distance with Tier 0 firms in

Table 8.4 Geographical distances between transaction partners

	No. of obs.	Mean	Std. dev.	p50	p25	p75
All	34,61,510	169.64	271.73	28.84	6.74	231.67
Transact. partn. ≥ 100	29,02,788	259.50	310.41	118.61	15.82	456.17
Transact. partn. ≥ 500	14,02,642	310.53	324.23	207.27	25.43	486.45
Transact. partn. ≥ 1000	9,38,824	329.38	328.93	257.24	28.98	490.85

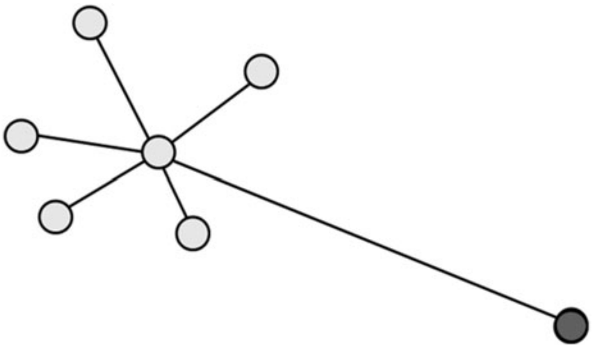


Fig. 8.5 Schematic representation of a hub firm

the affected areas, while at the same time having a large number of links with firms in their region. As a result, there are a large number of Tier 2 firms despite the relatively small number of Tier 1 firms in unaffected areas. Figure 8.5 presents a graphic depiction of the pattern, where the length of a line represents the distance between firms. The gray circle represents a firm in the affected areas, while the white circles represent firms in an unaffected area.

Lastly, let us examine how firms would be linked to those in the earthquake affected areas if there were no hub firms. In other words, let us consider a counterfactual network in which hub firms are removed from the actual network linking firms. Employing the same definition of hub firms as above, the share of hub firms with at least 1,000 relationships is 0.06 %, that of hub firms with at least 500 is 0.53 %, and that of hub firms with at least 100 is 1.34 %. Table 8.5 shows the pattern of firm links using the third definition (i.e., hub firms with at least 100 relationships) for the counterfactual. The table shows that the share of Tier 2 and Tier 3 firms is substantially lower than in Table 8.1. Moreover, even when longer path lengths are considered (i.e., Tier 4 and 5 firms), the share of firms linked to those in affected areas is still smaller than in Table 8.1. Thus, even though hub firms made up only 1.34 % of all firms, they play a key role in linking firms across longer distances. This, in turn, implies that hub firms played a key part in the spread of the economic impact of the earthquake.

Table 8.5 Cumulative percentage of firms in the network without hub firms

	Tier 0 (%)	Tier 1 (%)	Tier 2 (%)	Tier 3 (%)	Tier 4 (%)	Tier 5 (%)
Total	2	4	20	56	79	84
Hokkaido	0	1	20	65	86	88
Tohoku	17	27	57	81	89	91
Kanto	0	2	20	55	75	80
Chubu	0	1	14	54	82	88
Kinki	0	1	14	50	75	81
Chugoku/Shikoku	0	0	11	49	80	86
Kyushu	0	0	10	45	79	87

8.5 Conclusion

This chapter examined how the impact of the Great East Japan Earthquake spread geographically to physically unaffected areas through interfirm transaction networks. The main purpose was to address the conflicting implications regarding the geographical impact of the earthquake of the findings of previous studies on the structure of networks in Japan, namely that they are characterized by geographical proximity on the one hand and small-world structure on the other. The first, i.e., geographical proximity, implies that the economic impact of the earthquake should have remained locally confined, while small-world structure implies that it should have spread widely, as it in fact did.

Specifically, using interfirm transaction data from approximately 800,000 firms, the links between firms in the unaffected areas and those in the affected areas were examined. It was found that only 3 % of firm in the affected areas had direct transaction relationships with firms in the affected areas. On the other hand, once indirect links were taken into account, this share rises to 40–60 %. Furthermore, the geographical proximity of indirect links for firms throughout Japan was examined, showing that the median distance between indirect transaction partners is 255 km. This is much larger than the distance between direct partners (29 km) and similar to that between randomly chosen partners (416 km).

The fact that the share of direct transaction partners in the unaffected areas is very small is consistent with the fact that direct transactions tend to be localized. On the other hand, the analysis of indirect links showed that these are not localized. This result is consistent with the fact that networks are characterized by a small-world structure. Finally, by examining the paths linking indirect transaction partners, it was found that the observed differences between direct and indirect links regarding geographical proximity were due to regional hub firms which conduct transactions with firms outside their region.

These findings highlight that it is important to take into account that firms are closely linked with each other even if they are geographically distant and that regional hub firms play a key role not only in potentially spreading localized shocks but, by extension, also in preventing them from spreading.

Appendix

The tables in this appendix provide a more detailed breakdown of the share of firms linked to firms in the affected areas. Specifically, Tables 8.6 and 8.7 present the cumulative percentages of firms by industry, where Table 8.7 focusing on manufacturing industries presents those by more detailed category (two digit industry categories). Table 8.8 shows the cumulative percentages of firms by prefecture.

Table 8.6 Cumulative percentage of firms by industry

	Tier 0 (%)	Tier 1 (%)	Tier 2 (%)	Tier 3 (%)	Tier 4 (%)	Tier 5 (%)
Agriculture	2	5	40	84	96	97
Forestry	3	7	47	84	94	95
Fishery	10	16	57	91	96	96
Mining	3	7	54	92	96	97
Construction	2	5	44	88	97	98
Manufacturing	1	5	64	95	98	98
Water supply, gas, heating, electricity	3	7	70	86	88	88
Information and communication	1	2	68	90	93	93
Transportation	2	6	63	92	94	94
Wholesale and retail trade	2	6	70	95	98	98
Finance and insurance	1	3	50	72	76	76
Real estate	1	2	35	71	83	85
Dining and lodging	2	4	45	86	91	92
Medical, health care and welfare	2	2	45	76	78	78
Education, learning support	2	3	33	73	78	79
Compound services	3	7	48	88	93	94
Services (not elsewhere classified)	2	4	56	87	91	92
Public service (not elsewhere classified)	7	13	40	73	80	80

Table 8.7 Cumulative percentages of manufacturing firms

	Tier 0 (%)	Tier 1 (%)	Tier 2 (%)	Tier 3 (%)	Tier 4 (%)	Tier 5 (%)
Food	3	9	66	95	98	98
Beverages, tobacco, and feed	2	6	65	95	98	98
Textiles (excluding other textile products, clothing)	0	2	54	92	98	98
Apparel and other textile products	1	4	60	91	96	97
Lumber and wood manufacturing (except furniture)	3	10	58	91	99	99
Furniture and equipment manufacturing	1	4	56	94	99	99
Pulp, paper, and paper products	1	5	70	97	99	99
Printing and allied industries	1	3	61	95	98	98
Chemicals	1	8	72	94	96	97
Petroleum products and coal products	3	11	76	94	97	97
Plastic products (excluding those listed elsewhere)	1	4	66	96	98	99
Rubber products	1	5	61	95	98	98
Leather, fur, and their products	0	3	50	90	98	98
Ceramic, stone, and clay products	2	7	58	94	98	98
Steel	1	5	71	97	99	99
Non-ferrous metals	1	5	69	97	99	99
Fabricated metal products	1	4	63	97	99	99
General machinery, equipment, and supplies	1	5	68	97	99	99
Electrical machinery, equipment, and supplies	1	6	74	97	98	99
Information and communication electronics equipment	2	9	76	96	97	97
Electronic parts and devices	2	6	70	96	97	98
Transportation equipment	1	4	62	95	98	98
Precision machinery, equipment, and supplies	1	6	69	95	98	98
Other manufacturing	1	4	56	92	97	98

Table 8.8 Cumulative percentages of firms by prefecture

	Tier 0 (%)	Tier 1 (%)	Tier 2 (%)	Tier 3 (%)	Tier 4 (%)	Tier 5 (%)
Total	2	5	57	90	96	97
Hokkaido	0	2	60	96	99	99
Aomori	25	50	92	98	99	99
Iwate	22	46	94	98	99	99
Miyagi	48	77	96	98	99	99
Akita	0	16	82	98	99	99
Yamagata	0	15	80	97	98	98
Fukushima	23	40	89	98	99	99
Ibaragi	0	4	63	92	97	98
Tochigi	0	2	60	95	98	99
Gunma	0	1	49	89	97	97
Saitama	0	2	54	87	93	93
Chiba	0	2	58	88	93	93
Tokyo	0	4	63	88	92	93
Kanagawa	0	2	56	88	95	96
Niigata	0	3	57	93	97	98
Toyama	0	1	44	89	97	98
Ishikawa	0	1	45	88	96	97
Fukui	0	0	44	90	98	98
Yamanashi	0	0	48	94	98	99
Nagano	0	1	53	93	98	98
Gifu	0	1	47	85	94	95
Shizuoka	0	1	51	90	97	97
Aichi	0	1	56	93	98	98
Mie	0	0	53	91	97	98
Shiga	0	1	47	80	90	91
Kyoto	0	1	47	85	94	96
Osaka	0	2	60	90	95	96
Hyogo	0	1	53	86	94	95
Nara	0	0	50	86	95	96
Wakayama	0	0	45	86	97	98
Tottori	0	0	38	87	96	98
Shimane	0	0	39	85	95	96

Table 8.8 (continued)

	Tier 0 (%)	Tier 1 (%)	Tier 2 (%)	Tier 3 (%)	Tier 4 (%)	Tier 5 (%)
Okayama	0	0	44	90	97	98
Hiroshima	0	1	50	90	96	96
Yamaguchi	0	1	54	89	94	95
Tokushima	0	0	41	87	96	96
Kagawa	0	1	52	94	98	99
Ehime	0	0	52	94	99	99
Kochi	0	0	40	88	97	97
Fukuoka	0	0	52	89	96	97
Saga	0	0	41	84	94	95
Nagasaki	0	0	43	88	97	98
Kumamoto	0	0	36	84	95	96
Oita	0	0	41	91	98	98
Miyazaki	0	0	38	90	98	98
Kagoshima	0	0	42	90	97	98
Okinawa	0	0	29	88	98	99

References

Albert, R., & Barabási A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*, 47–97.

Duranton, G., & Puga D. (2004). Micro-foundations of urban agglomeration economies. In V. Henderson & J. Thisse (Eds.), *Handbook of regional and urban economics 4* (pp. 2063–2117). North-Holland, Amsterdam.

Ellison, G., Glaeser, E., & Kerr, W. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review, 100*(3), 1195–1213.

Gabaix, X. (2009). Power law in economics and finance. *Annual Review of Economics, 1*, 255–293.

Nakajima, K., Saito Y. U., & Uesugi, I. (2012). Localization of interfirm transaction relationships and industry agglomeration. Discussion papers 12-E-023, Research Institute of Economy, Trade and Industry.

Nakajima, K., Saito Y. U., & Uesugi, I. (2013). Role of inter-firm transactions on industrial agglomeration: Evidence from Japanese firm-level data. Discussion papers 13-E-021, Research Institute of Economy, Trade and Industry.

Ohnishi, T., Takayasu, H., Takayasu, M. (2010). Network Motifs in Inter-firm Net-work. *Journal of Economic Interaction and Coordination, 5*(2), 171–180.

Saito, Y. U., & Watanabe T. (2007). Do larger firms have more interfirm relationships?. *Physica A, 383*, 158–163.

Rosenthal, S., & Strange, W. (2001). The determinants of agglomeration. *Journal of Urban Economics, 50*, 191–229.

Yukiko Umeno Saito is a senior fellow at the Research Institute of Economy, Trade and Industry, Japan since 2012, and was a researcher at the Fujitsu Research Institute, Japan from 2002 to 2012. She has participated in several research projects by institutions, National Institute of Science and Technology Policy, National Institute for Research Advancement, and so forth. She received a Ph.D. in physics from University of Tokyo in 2002. She was a research follow of Japan Society for the Promotion of Science from 1999 to 2002. Her research field is industrial organization and spatial economics. She is interested in how firms construct inter-firm relationships and how it affects firm's performance, especially propagation through the relations.

Part III
Bank-Firm Relationships and Firm Dynamics

Chapter 9
Bank-firm Relationships: A Review of the Implications for Firms and Banks in Normal and Crisis Times

Hans Degryse, Vasso Ioannidou and Steven Ongena

Abstract Banks are important providers of external finance to firms. In order to solve asymmetric information problems, firms and banks often engage in bank-firm relationships. Relationship banking occurs when a bank and a borrower enter multiple mutual interactions and both parties invest in obtaining some counterparty specific information, binding bank and firm, to a certain degree, to each other. This chapter starts with a discussion of reasons for having exclusive versus non-exclusive relationships. It provides a concise overview on the determinants of the number and intensity of bank-firm relationships, and reviews how relationship banking generates costs and benefits for both banks and firms. We show that on average bank-firm relationships generate value for both. The costs and benefits of bank-firm relationships, however, vary substantially with whether an economy is in normal or crisis times.

Keywords Relationship banking · Non-exclusivity · Financial crisis

9.1 Introduction

The financial sector has been under severe stress in the last couple of years. The 2007–2009 financial crisis revealed that banks were exposed to risks which were far removed from their core business. Banks and financial institutions from around

---

The first and second author gratefully acknowledge support from FWO-Flanders under contract G.0719.13.

H. Degryse (✉)
KU Leuven and CEPR,
Naamsestraat 69, 3000 Leuven, Belgium
e-mail: Hans.Degryse@kuleuven.be

V. Ioannidou
Lancaster University, LA1 4YX Lancaster, UK
e-mail: v.ioannidou@lancaster.ac.uk

S. Ongena
University of Zurich, SFI and CEPR,
Plattenstrasse 32, 8032 Zurich, Switzerland
e-mail: steven.ongena@bf.uzh.ch

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_9

the world had to write off trillions of dollars in losses. The financial crisis had a profound impact on the real sector with recessions in many countries. Researchers, policy makers, and regulators have suggested that banks should move away from the originate-to-distribute model and go back to basics by focusing on their core functions, i.e., the gathering of deposits and the origination of (relationship) loans to borrowers where banks are perceived to be specialist institutions and to subsequently keep those loans on their balance sheets.

The recent financial crisis induced new regulation such as the new Basle III capital adequacy rules and the Dodd-Frank Act in the United States, for example. Supervision in many countries has been put under scrutiny and is being reorganized. An interesting case is the creation of the European banking union.

In this paper we briefly review the literature on bank-firm relationships and how bank-firm relationships fare in normal and crisis times. We focus on relationship banking which is at the heart of the back-to-the-basics argument. Relationship banking implies investment in counterparty-specific information, binding bank and firm, to a certain degree, to each other. We summarize how relationship banking generates costs and benefits for both banks and firms, but argue that on average it generates value for both of them. The impact, however, depends largely on whether an economy is in normal or crisis times.

The remainder of this chapter is organized as follows. Section 9.2 starts our review by discussing reasons for why banks and firms may engage into exclusive versus non-exclusive credit. Section 9.3 then discusses the costs and benefits of intimate borrowing relationships (relationship lending) both in normal and crisis times. Section 9.4 reviews the role of bank-firm relationships in the transmission of monetary policy. Section 9.5 concludes.

9.2 Exclusive Versus Non-Exclusive Lending

An important characteristic of financial contracts is that they are non-exclusive. In credit markets, for example, borrowers cannot credibly commit to take loans from at most one bank and banks cannot completely prevent borrowers from taking credit from other financial institutions, issue public debt, or tap abnormal amounts of trade credit. This is because contracts cannot be made fully contingent on loans from other creditors and in particular on future creditors who have not yet lent to the borrower. Such loans, however, could adversely affect a borrower's probability of repayment by exacerbating moral hazard incentives and incentives for strategic default (e.g., Bizer and DeMarzo 1992; Parlour and Rajan 2001). The prospect of such loans is expected to worsen the borrower's access and terms of credit. When non-exclusivity is pervasive and cannot be contained, it could also lead to overborrowing, high rates of default, credit rationing, and market freezes (see e.g., Bennardo et al. (2014)).

Non-exclusivity is not only important for loan contracts between firms and creditors but it also generalizes to other financial contracts. Non-exclusivity in the credit derivatives market, for example, has played a central role in the financial crisis of

2007–2009. Acharya and Bisin (2014) argue that the non-exclusivity of financial contracts coupled with the opacity of the over-the-counter (OTC) markets—where credit default swaps (CDS) trade—played a central role in the current financial crisis by creating severe counterparty risk externalities. The risk that a party—in this case the seller of a CDS—might not be able to fulfil its future obligations depends largely on other, often subsequent, exposures. In a theoretical model, Acharya and Bisin (2014) show that more transparency on counterparty risk exposures in the OTC market could have helped the contracting parties to internalize the externalities. These insights are in line with parallel theoretical work on the role of the institutional framework on credit markets. Collateral and credit registries, for example, could help banks protect their claims and thus dampen the impact of non-exclusivity on credit availability. Collateral, whose effective use is facilitated by a collateral registry, could mitigate moral hazard incentives and incentives for strategic default (Holmstrom and Tirole 1997; Parlour and Rajan 2001). Credit registries, for example, could in some cases allow lenders to effectively employ ex-post punishment to enforce exclusivity or mitigate the resulting externalities by conditioning their terms on loans from others (Bennardo et al. 2014).

Degryse et al. (2013) employ a unique dataset containing information on a creditor's internal limit to the borrower both before and after a non-exclusivity event realizes. The internal limit indicates the maximum amount this creditor is willing to lend to a borrower; it represents the amount for which the bank's loan supply becomes vertical. Changes in the internal limit represent changes in loan supply. Hence, using this information, they investigate how a creditor's willingness to lend reacts after a firm with whom it held an exclusive relationship acquires loans from other creditors, which they refer to as "outside loans". The data set they employ is from Sweden implying that the empirical analysis takes place in a setting where individual trades with other creditors can be observed and contractual features, such as collateral, can be employed more effectively.

Degryse et al. (2013) find evidence consistent with the theories on contractual externalities. In particular, they find that when a previously exclusive firm, obtains a loan from another bank, the firm's initial bank decreases its internal limit to the firm and it decreases it more the larger the size of the outside loans. In terms of magnitudes, they find that a $ 1 from another bank leads to a decrease in the initial bank's willingness to lend by 34–50 cents. Furthermore, consistent with the theoretical literature on contractual externalities, they also find that the initial bank's willingness to lend does not change when its existing and future loans are protected from the increased risk of default. In particular, the authors find that an outside loan does not trigger any change in the initial bank's willingness to lend if its existing and future loans retain seniority over the outside loans *and* their claims are secured with assets whose value is high and stable over time.

Engaging multiple banks may also allow firms to solve some problems they may have with their so-called relationship banks—see more on this in Section 9.3. For example, it may allow firms to reduce hold-up problems (Fischer 1990; Sharpe 1990; Rajan 1992; von Thadden 2004) and dampen shocks impacting the liquidity of their banks (Detragiache et al. 2000). The willingness of another bank to extend credit to a firm may also be perceived as a positive signal about its quality (James 1987) and thus boost the initial bank's willingness to lend to the firm, particularly when the initial bank is relatively uninformed. Findings in Degryse et al. (2013) support this hypothesis for small and young firms.

Several papers have investigated the reasons and the impact of establishing single versus multiple bank relationships. Ongena and Smith (2000), for example, find that older and larger firms and firms in countries with a lower degree of judicial efficiency are more likely to maintain multiple relationships (for an extensive overview of empirical studies in this literature see e.g., Degryse et al. (2009)). Some papers also find that firms that borrow from multiple banks are of lower quality (see, for example, Petersen and Rajan (1994)). Farinha and Santos (2002) follow the debt share of firms after initiating multiple relationships. They find that the bank with which the firm had an exclusive relationship only provided about half of the firm's bank debt after 3 years.

Braggion and Ongena (2012) relate firms' leverage and debt financing to the secular behavior of bank-firm relationships. They argue that the penchant of corporates to borrow from multiple banks operating in a competitive banking market may be an important driver of corporate leveraging. Competing banks may fail, for example, to fully internalize the consequences of future corporate indebtedness depending upon the institutional environment especially when vying for market share. As a consequence banks may "overlend." Braggion and Ongena (2012) study data for a large sample of UK firms between 1896 and 1986 and they document that with the onset of banking sector deregulation in 1970 a subsequent and remarkable shift from bilateral to multilateral relationship banking took place. They then relate such a shift to firms' use of debt finance and its effect on leverage ratios.

In the next section, we review the benefits and costs of establishing bank-firm relationships in normal and crisis times.

9.3 Relationship Banking: Costs and Benefits for Firms and Banks in Normal and Crisis Times

Banks and firms often engage in long-term relationships. These may be beneficial for both banks and firms. Its magnitude however depends on whether we face normal times or crisis times. Boot (2000) defines relationship banking as "the provision of financial services by a financial intermediary that invests in obtaining customer-specific information, often proprietary in nature and evaluates the profitability of these investments through multiple interactions with the same customer over time and/or across products" (p. 10). It is through the temporal progression of a relationship that a bank can learn more than other financiers about a firm's ability to meet future obligations, either through the monitoring of debt covenants and payment history or through other services offered to the firm by the bank. For example, the bank may piece together an accurate picture of the firm by looking at past activities on the firm's checking account which may be helpful in more accurately assessing default and give the ‘inside bank’ an informational advantage over outside banks (Nakamura 1993; Mester et al. 2007; Norden and Weber 2010). It is this informational asymmetry between the inside bank and other ‘outside’ banks, which gives the inside bank a competitive edge and binds firms to banks implying an almost assures continued interaction between the bank and its high-quality borrowers (Fischer 1990; Sharpe 1990; Rajan 1992; von Thadden 2004; Ioannidou and Ongena 2010).

Bank-firm relationships can be found in most advanced economies but the duration and importance tends to vary considerably across countries as well as across firms. Degryse et al. (2009) review several characteristics of bank-firm relationships such as the duration and number of relationships held in several countries. For example, they show that the average duration of bank-firm ranges from 8 years in Belgium (Degryse and Van Cayseele 2000) and 7–11 years in the United States (Petersen and Rajan 1994) to 13 years in Germany (Harhoff and Körting 1998), 14 years in Italy (Angelini et al. 1998), and 15–18 years in Norway (Ongena and Smith 2001). Bank-firm relationships are also important in Japan with average durations of over 20 years (e.g., Horiuchi et al. 1988; Uchida et al. 2008).

We now investigate how relationship banking impacts credit allocation and how this has benefits and costs to both banks and firms, providing us a picture on the value for both parties involved.

9.3.1 Bank Relationships: Benefits and Costs for the Firm During Normal and Crisis Times

Firms may benefit from the availability, flexibility, control, reputation, and confidentiality embedded in a bank relationship. First, there is quite a large empirical literature showing that a credit relationship increases access to capital, possibly at a lower cost and/or with less collateral. These beneficial credit contract characteristics improve credit allocation and spur firm growth. In addition to increased availability, a credit relationship may foster ex-ante flexibility in writing loan contracts and allow a firm to fulfill its more complex and non-standard credit needs (Boot and Thakor 1994; von Thadden 1995). For a firm experiencing difficulties meeting contracted loan payments, a bank can smooth interest rates and reschedule capital payments through, for example, overdraft facilities and renegotiation (Chemmanur and Fulghieri 1994). But the bank may also either accommodate the firm with new lending or refuse future lending, conditional on actions taken by the firm during and after the distress period. Thus, banks may have the ability to exert control over the management of firm assets, which may induce managers to take optimal decisions (Rajan 1992). If repeated lending from a reputable financial institution provides credible certification and control of managers’ actions, a credit relationship may also bolster the firm’s reputation. Immaculate standing may facilitate current and future funding from both shareholders and alternative outside sources (Diamond 1991). The confidentiality of a bank relationship may also further facilitate screening and monitoring (Campbell 1979), may prevent leakage of proprietary information to product market competitors (Yosha 1995; von Rheinbaben and Ruckes 2004; Degryse and Ongena 2001), and may encourage investment in Research and Development (Bhattacharya and Chiesa 1995).

In crisis times, firms may prefer to solve their expected financial problems privately in a credit relationship, rather than damaging their reputation on the financial markets. Theory suggests that relationships are valuable during crisis times, as banks are able to smoothen out shocks to firms. Evidence for this is provided by Jiangli et al. (2008) where relationships help in some countries during crisis times. However, when the crisis is really systemic, this seems to be less the case, a topic on which we turn to below.

The ability for a bank to privately observe proprietary information and maintain a close relationship with its customer can also impose *costs* on the customer. For example, an inside bank has the ability to offer only above-cost loans to its best customers and hold-up customers from receiving competitive financing elsewhere. The inside bank gains this monopoly power through its informational advantage over competitors. A high-quality firm that tries to switch to a competing uninformed bank gets pooled with low-quality firms and is offered an even worse, breakeven interest rate (Sharpe 1990; Rajan 1992; von Thadden 2004). The costs arising from hold-up problems may also be tempered by the bank's desire to acquire a reputation for refraining from extracting monopoly hold-up rents (Sharpe 1990) or for financing productive firms by making more efficient continuation decisions in renegotiation (Chemmanur and Fulghieri 1994). Ioannidou and Ongena (2010) provide convincing direct evidence on the existence of hold-up problems and associated costs for customers. They do this employing matching techniques to make sure that firms are similar from the point of view of the inside bank. They find that switchers to other banks obtain loans at about 80 basis points lower interest rates than similar non-switching firms staying with the inside bank (see also Degryse et al. (2009), their Chap. 4, for a review on indirect evidence for the existence of hold-up problems).

Though informationally transparent firms are less likely to be affected by this hold-up problem, one solution for more opaque firms is suggested by von Thadden (1992). Establishing multiple firm-bank relationships can create competition among banks and can therefore limit each individual bank's rent extraction ability. Other papers in the literature investigate the optimal number of creditors. Bolton and Scharfstein (1996) and Bris and Welch (2005), for example, explore the impact of the debt structure on the efficiency of the renegotiation that may take place in the case of firm default. Detragiache et al. (2000) model how firms may seek to diversify bank liquidity risk by engaging multiple financiers. Another strand in the literature explores the banks' monitoring incentives. Carletti (2004), for example, argues that firms may benefit from borrowing from two banks to mitigate the excessive monitoring that takes place when only one bank is engaged (see also Carletti et al. 2007).

Furthermore, there may be supply induced credit availability effects stemming from banks during a financial crisis that have an economically significant impact on firms. Ivashina and Scharfstein (2010), for example, show that firms run on their credit lines during the recent financial crisis and that they have difficulties in renewing these lines. Strong bank relationships do not seem to help when the banking crisis is systemic. Recent empirical evidence by Carvalho et al. (2010), for example, shows that firms with more intense relationships with banks that face larger shocks (announcements of bank asset write-downs) during the 2007–2008 financial crisis suffer greater equity valuation losses. These impacts are not offset by borrowers' access to public debt markets and the impact is largest for firms with the greatest information asymmetry problems. This somehow suggests that banks are able to smoothen out idiosyncratic shocks but may amplify systemic shocks. These results, however, seem not to necessarily carry over towards smaller borrowers. Puri et al. (2011), for example, employ loan application data for retail loans at German savings banks in the period 2006–2008. They investigate whether savings banks that are exposed to shocks from Landesbanken (whom they own) stemming from the US, behave differently than non-exposed savings banks, i.e., who own Landesbanken without exposure to the US financial crisis. They find evidence for a supply-side effect in that the affected banks reject substantially more loan applications than non-affected banks. Furthermore, bank relationships mitigate supply side effects as firms with longer relationships are less likely to be rejected even when their savings bank is exposed to a financial shock.

Finally, important to notice is also that banking crises are not exogenous phenomena, but regularly come on the heels of periods of strong credit growth (Kindleberger 1978; Schularick and Taylor 2012; Gourinchas and Obstfeld 2012). As discussed above, Braggion and Ongena (2012) relate firms' leverage and debt financing to the secular behavior of bank-firm relationships. They argue that the penchant of corporates to borrow from multiple banks operating in a competitive banking market may be an important driver of corporate leveraging.

9.3.2 Bank Relationships: Benefits and Costs for the Bank

In the previous subsection, we discussed the benefits and costs for firms. Often what is perceived as costs for firms are the benefits for banks and vice versa. Banks invest in bank-firm relationships. This relationship capital may affect the way banks allocate credit across groups of borrowers. Banks can extract higher profits from captured and locked-in (opaque) borrowers than from borrowers that have ready access to other financing alternatives. As explained above, banks are able to charge relatively higher interest rates to opaque borrowers once the bank has sunk its initial relationship cost. This may still be beneficial for the firm as otherwise the firm would need to incur the sunk cost of switching to another bank.

Kim et al. (2003) investigate the value to the bank of a locked in customer employing Norwegian data. They find that the marginal value of lock-in for a bank is 0.16; that is, 16 % of the customer's added value is attributed to the lock-in phenomenon generated by switching costs. They further find that the contribution of locked-in customers to banks' value decrease as the size of bank increases. Specifically, the contribution of locked-in customers to banks' value ranges from a low of 13 % for the very large banks to 30 % for the group including also the smaller ones when the market is defined according to branch-network size.

Summarizing, it seems that both partners to a lending relationship derive positive value from this engagement. Firms derive accreditation and the economic value emanating from it, as highlighted in the work by James (1987), who finds that bank loan announcements are associated with positive and statistically significant stock price reactions (shown to equal about 200 basis points in a 2-day window), while announcements of privately placed and public issues of debt experience zero or negative stock price reactions. The positive stock-price reaction supports the Fama (1985) argument that a bank loan provides accreditation for a firm's ability to generate a certain level of cash flows in the future. This contribution to borrowers is accompanied by gains to the bank emanating from borrowers' captivity generated by high switching costs. Significant switching costs confer positive marginal value to the bank from a borrower's lock-in and can contribute to about 13–30 % of the bank's value, depending upon its size.

9.4 The Impact of Monetary Policy and Bank-Firm Relationships

While little direct theoretical modeling or empirical evidence exists on the impact of monetary or business cycle conditions on the duration, scope or multiplicity of bank-firm relationships, there is plenty of thinking and evidence on the external finance premium to which relationship formation and characteristics may respond.

The external finance premium in lending depends inversely on the borrowers' net worth (see Freixas and Rochet 2008). When borrowers have little wealth at stake, the potential divergence of interests between the borrower and the suppliers of external funds is larger, increasing agency costs. In equilibrium, lenders must be compensated. As borrower net worth is pro-cyclical (because profits and asset prices are pro-cyclical), the external finance premium is countercyclical, amplifying the changes in credit availability (Bernanke et al. 1999; Matsuyama 2007). In Holmstrom and Tirole (1997) the agency problems depend on the capital-to-total-assets ratio, in Bernanke et al. (1999) net worth is also associated with the liquidity of the assets. Since banks not only face agency problems with their borrowers, but banks themselves are also borrowing funds from their depositors and other financiers, bank net worth may determine their own agency costs of borrowing (Bernanke 2007; Boivin et al. 2011; Gertler and Kiyotaki 2011).

Translating this to relationship formation, persistence and intensity one can argue that in times of lax monetary conditions and during boom times, and in times of low overall uncertainty about economic conditions, borrower net worth is high and hence it is easy for firms to find new banks. Many new relationships will be formed. In contrast when monetary conditions have been tight for a while, the economy is going in a through or during crisis times forming a new relationship will be more difficult and expensive. Existing relationships then "deliver" so to speak, in terms of guaranteeing credit continues to flow to the firm. And relation lenders can therefore charge a higher rate during normal times (Bolton et al. 2013), to subsidize the firm during bad times.

There may also be compositional effects, in the sense that banks have a different willingness to different types of borrowers depending on the monetary and business cycle conditions that prevail. There is evidence in the literature that suggests a correlation between the monetary policy rate and financiers' risk-taking. Adrian and Shin (2011) in their discussion of the risk-taking channel of monetary policy document correlations between short-term interest rates and bank risk-taking (see also De Nicolò et al. (2010)). Den Haan et al. (2007) suggest that high short-term rates could imply a decline in bank risk-taking with US data, and Maddaloni and Peydró (2011) with Euro area data. And Gertler and Gilchrist (1994) show that contractionary monetary policy results in less bank lending to small firms, findings that are consistent not only with the firm balance-sheet channel, but also with possibly less bank risk-taking.

Recently Jiménez et al. (2013) and Ioannidou et al. (2014) further investigate the impact of monetary policy on banks' risk-taking. Their estimates suggest that a lower monetary policy rate spurs bank risk-taking and hence that monetary policy affects the composition of the supply of credit beyond the well-documented effects of both the bank- and firm balance-sheet channels and changes in the demand for credit. And ongoing empirical work documents the robust existence and potency of such a bank risk-taking channel across countries and time periods, e.g., for the US (Altunbas et al. 2010; Delis et al. 2011; Paligorova and Santos 2012), Austria (Gaggl and Valderrama 2010), Colombia (López et al. 2010a, b), the Czech Republic (Geršl et al. 2012), and Sweden (Apel and Claussen 2012).

What does this all implies for bank-firm relationships? If banks shift their lending during expansionary monetary conditions towards more risky firms (that potentially pay to start "a relationship"), this will imply that in contractionary times banks may be saddled with a portfolio of firms that are potentially bad risks. Some of these firms will disappear, but for other firms the banks will have deliver its part of the intertemporal bargain.

9.5 Conclusion

The 2007–2008 financial crisis has put the worldwide financial sector under severe stress. Researchers, policymakers, and regulators argued that banks should move back to their core intermediation function and move away from the originate-to-distribute model.

In this paper we investigate the roots of this argumentation and summarize how relationship banking generates both costs and benefits for both firms and banks, but that on average it generates value for both of them. The value of relationships, however, depends on whether an economy is in normal or crisis times. While banks are able to provide insurance to relationship borrowers hit by idiosyncratic shocks in normal times, it is less clear whether banks are able to fulfill this role in crisis times.

References

Acharya, V. V., & Bisin, A. (2014). Counterparty risk externality: Centralized versus over-the-counter markets. *Journal of Economic Theory, 149*, 153–182.

Adrian, T., & Shin, H. S. (2011). Financial intermediaries and monetary economics. In: B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics 3A*. Elsevier: New York.

Altunbas, Y., Gambacorta, L., & Marquez-Ibañez, D. (2010). *Does monetary policy affect bank risk-taking*? Working Paper, Bank for International Settlements.

Angelini, P., Di Salvo, R., & Ferri, G. (1998). Availability and cost of credit for small businesses: Customer relationships and credit cooperatives. *Journal of Banking and Finance, 22*, 925–954.

Apel, M., & Claussen, C. A. (2012). Monetary policy, interest rates and risk-taking. *Sveriges Riksbank Economic Review, 2,* 68–83.

Bennardo, A., Pagano, M., & Piccolo, S. (2014). Multiple-bank lending, creditor rights and information sharing. *Review of Finance*. (forthcoming).

Bernanke, B. S. (2007). The financial accelerator and the credit channel. Remarks-credit channel of monetary policy in the twenty-first century, Board of Governors of the US Federal Reserve System.

Bernanke, B. S., Gertler, M., & Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. In T. John & W. Michael (Eds.), *Handbook of macroeconomics*. Amsterdam: Elsevier.

Bhattacharya, S., & Chiesa, G. (1995). Proprietary information, financial intermediation, and research incentives. *Journal of Financial Intermediation, 4*, 328–357.

Bizer, D. S., & DeMarzo, M. (1992). Sequential banking. *Journal of Political Economy, 100*, 41–61.

Boivin, J., Kiley, M., & Mishkin, F. S. (2011). How has the monetary transmission mechanism evolved over time? In B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics 3A*. New York: Elsevier.

Bolton, P., & Scharfstein, D. S. (1996). Optimal debt structure and the number of creditors. *Journal of Political Economy, 104*, 1–25.

Bolton, P., Freixas, X., Gambacorta, L., & Mistrulli, P. E. (2013). *Relationship and transaction lending in a crisis*. Bank for International Settlements, working paper 417, 68 pp.

Boot, A. W. A. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation, 9*, 3–25.

Boot, A. W. A., & Thakor, A. V. (1994). Moral hazard and secured lending in an infinitely repeated credit market game. *International Economic Review, 35*, 899–920.

Braggion, F., & Ongena, S. (2012). *A century of firm–bank relationships: Did banking sector deregulation spur firms to add banks and borrow more*? Mimeo, Tilburg: The Netherlands.

Bris, A., & Welch, I. (2005). The optimal concentration of creditors. *Journal of Finance, 60*, 2193–2212.

Campbell, T. S. (1979). Optimal investment financing decisions and the value of confidentiality. *Journal of Financial and Quantitative Analysis, 14*, 232–257.

Carletti, E. (2004). The structure of bank relationships, endogenous monitoring, and loan rates. *Journal of Financial Intermediation, 13*, 58–86.

Carletti, E., Cerasi, V., & Daltung, S. (2007). Multiple-bank lending: Diversification and free-riding in monitoring. *Journal of Financial Intermediation, 16*, 425–451.

Carvalho, D., Ferreira, M. A., & Matos, P. (2010). *Lending relationships and the effect of bank distress: Evidence from the 2007–2008 financial crisis*. Mimeo, Nova: Lisbon.

Chemmanur, T. J., & Fulghieri, P. (1994). Reputation, renegotiation and the choice between bank loans and publicly traded debt. *Review of Financial Studies, 7*, 475–506.

De Nicolò, G., Dell'Ariccia, G., Laeven, L., & Valencia, F. (2010). *Monetary policy and bank risk-taking*. Mimeo, International Monetary Fund, Washington DC.

Degryse, H., & Ongena, S. (2001). Bank relationships and firm performance. *Financial Management, 30*, 9–34.

Degryse, H., & Van Cayseele, P. (2000). Relationship lending within a bank-based system: Evidence from European small business data. *Journal of Financial Intermediation, 9*, 90–109.
Degryse, H., Kim, M., & Ongena, S. (2009). *Microeconometrics of banking: Methods, applications and results*. Oxford University Press: Oxford
Degryse, H., Ioannidou, V., & von Schedvin, E. L. (2013). On the non-exclusivity of loan contracts: An empirical investigation. Discussion Paper, Centre for Economic Policy Research.
Delis, M. D., Hasan, I., & Mylonidis, N. (2011). *The risk-taking channel of monetary policy in the USA: Evidence from micro-level data*. Mimeo, Cass Business School.
Den Haan, W. J., Sumner, S., & Yamashiro, G. (2007). Bank loan portfolios and the monetary transmission mechanism. *Journal of Monetary Economics, 54*, 904–924.
Detragiache, E., Garella, P. G., & Guiso, L. (2000). Multiple versus single banking relationships: Theory and evidence. *Journal of Finance, 55*, 1133–1161.
Diamond, D. W. (1991). Monitoring and reputation: The choice between bank loans and privately placed debt. *Journal of Political Economy, 99*, 689–721.
Fama, E. F. (1985). What's different about banks? *Journal of Monetary Economics, 15*, 5–29.
Farinha, L. A., & Santos, J. A. C. (2002). Switching from single to multiple bank lending relationships: Determinants and implications. *Journal of Financial Intermediation, 11*, 124–151.
Fischer, K. (1990). *Hausbankbeziehungen als Instrument der Bindung zwischen Banken und Unternehmen-Eine Theoretische und Empirische Analyse*. PhD Dissertation, Universitat Bonn.
Freixas, X., & Rochet, J. C. (2008). *Microeconomics of banking*. Cambridge: MIT Press.
Gaggl, P., & Valderrama, M. T. (2010). Does a low interest rate environment affect risk taking in Austria? *Monetary Policy and the Economy, 4,* 32–48.
Geršl, A., Jakubík, P., Kowalczyk, D., Ongena, S., & Peydró, J.-L. (2012). *Monetary Conditions and Banks' Behaviour in the Czech Republic*. Working Paper, Czech National Bank.
Gertler, M., & Gilchrist, S. (1994). Monetary policy, business cycles, and the behavior of small manufacturing firms. *Quarterly Journal of Economics, 109*, 309–340.
Gertler, M., & Kiyotaki, N. (2011). Financial intermediation and credit policy in business cycle analysis. In B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics 3A*. New York: Elsevier.
Gourinchas, P.-O., & Obstfeld, M. (2012). Stories of the twentieth century for the twenty-first. *American Economic Journal: Macroeconomics, 4*, 226–265.
Harhoff, D., & Körting, T. (1998). Lending relationships in Germany—Empirical evidence from survey data. *Journal of Banking and Finance, 22*, 1317–1353.
Holmstrom, B., & Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics, 112*, 663–691.
Horiuchi, T., Packer, F., & Fukuda, S. (1988). What role has the 'Main Bank' played in Japan? *Journal of Japanese and International Economies, 2*, 159–180.
Ioannidou, V. P., & Ongena, S. (2010). "Time for a change": Loan conditions and bank behavior when firms switch banks. *Journal of Finance, 65*, 1847–1878.
Ioannidou, V. P., Ongena, S., & Peydró, J.-L. (2014). Monetary policy, risk-taking and pricing: Evidence from a Quasi-natural experiment. *Review of Finance*. (Forthcoming).
Ivashina, V., & Scharfstein, D. S. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics, 97*, 319–338.
James, C. M. (1987). Some evidence on the uniqueness of bank loans. *Journal of Financial Economics, 19*, 217–235.
Jiangli, W., Unal, H., & Yom, C. (2008). Relationship lending, accounting disclosure, and credit availability during the Asian financial crisis. *Journal of Money, Credit and Banking, 40*, 25–55.
Jiménez, G., Ongena, S., Peydró, J.-L., & Saurina, J. (2013). *Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?* Mimeo, Bank of Spain: Madrid.
Kim, M., Kliger, D., & Vale, B. (2003). Estimating switching costs: The case of banking. *Journal of Financial Intermediation, 12*, 25–56.

Kindleberger, C. P. (1978). *Manias, panics, and crashes: A history of financial crises*. New York: Basic Books.
López, M., Tenjo, F., & Zárate, H. (2010a). *The risk-taking channel and monetary transmission mechanism in Colombia*. Working Paper, Banco de la Republica Colombia.
López, M., Tenjo, F., & Zárate, H. (2010b). *The risk-taking channel in Colombia revisited*. Working Paper, Banco de la Republica Colombia.
Maddaloni, A., & Peydró, J.-L. (2011). Bank risk-taking, securitization, supervision, and low interest rates: Evidence from Euro-area and U.S. lending standards. *Review of Financial Studies, 24*, 2121–2165.
Matsuyama, K. (2007). Credit traps and credit cycles. *American Economic Review, 97*, 503–516.
Mester, L. J., Nakamura, L., & Renault, M. (2007). Transactions accounts and loan monitoring. *Review of Financial Studies, 20*, 529–556.
Nakamura, L. I. (1993). Monitoring loan quality via checking account analysis. *Journal of Retail Banking, 14*, 16–34.
Norden, L., & Weber, M. (2010). Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies, 23*, 3665–3699.
Ongena, S., & Smith, D. C. (2000). What determines the number of bank relationships? Cross-country evidence. *Journal of Financial Intermediation, 9*, 26–56.
Ongena, S., & Smith, D. C. (2001). The duration of bank relationships. *Journal of Financial Economics, 61*, 449–475.
Paligorova, T., & Santos, J. A. C. (2012). *Monetary policy and bank risk-taking: Evidence from the corporate loan market*. Mimeo, Bank of Canada: Ottawa
Parlour, C., & Rajan, U. (2001). Competition in loan contracts. *American Economic Review, 91*, 1311–1328.
Petersen, M. A., & Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance, 49*, 3–37.
Puri, M., Rocholl, J., & Steffen, S. (2011). Global retail lending in the aftermath of the US financial crisis: Distinguishing between supply and demand effects. *Journal of Financial Economics, 100*, 556–578.
Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance, 47*, 1367–1400.
Schularick, M., & Taylor, A. M. (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870–2008. *American Economic Review, 102*, 1029–1061.
Sharpe, S. A. (1990). Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships. *Journal of Finance, 45*, 1069–1087.
Uchida, H., Udell, G. F., & Watanabe, W. (2008). Bank size and lending relationships in Japan. *Journal of the Japanese and International Economies, 22*, 242–267.
von Rheinbaben, J., & Ruckes, M. (2004). The number and the closeness of bank relationships. *Journal of Banking and Finance, 28*, 1597–1615.
von Thadden, E. L. (1992). *The commitment of finance, duplicated monitoring and the investment horizon*. Working Paper in Financial Markets, Centre for Economic Policy Research.
von Thadden, E. L. (1995). Long-term contracts, short-term investment, and monitoring. *Review of Economic Studies, 62*, 557–575.
von Thadden, E. L. (2004). Asymmetric information, bank lending, and implicit contracts: The winner's curse. *Finance Research Letters, 1*, 11–23.
Yosha, O. (1995). Information disclosure costs and the choice of financing source. *Journal of Financial Intermediation, 4*, 3–20.

Hans Degryse is Professor of Finance at the Department of Accountancy, Finance and Insurance of the KU Leuven. He is a research fellow at the CEPR, CESIfo, the European Banking Center (EBC), and TILEC, and is member of the academic council of EBC. Before joining Leuven in 2012, he was professor of finance at Tilburg University. His research focuses on financial intermediation,

including theoretical and empirical banking as well as market microstructure. He has published in many journals including the *American Economic Review, Journal of Finance, Journal of Financial Economics, Review of Financial Studies, Journal of Financial Intermediation*, and the *Economic Journal*, and it has been presented in leading international conferences such as the *American Finance Association*, the *Western Finance Association*, the *European Finance Association*, and the *Financial Intermediation Research Society*. He co-authored, with Moshe Kim and Steven Ongena, the graduate textbook *Microeconometrics of Banking: Methods, Applications and Results.*

Vasso Ioannidou is a Professor of Finance at the Department of Accounting and Finance at Lancaster University and a research fellow and board member of the European Banking Center (EBC). Prior to joining Lancaster in 2013, she was a Professor at the Department of Finance at Tilburg University. Her research covers topics related to the transmission mechanism of monetary policy, prudential regulation and supervision of banks, and topics related to the effects of asymmetric information and institutions on bank-firm lending relationships and access to finance. Her work has been published in leading and high quality academic journals such as the *Journal of Finance*, the *Journal of Financial Economics*, the *Journal of Financial Intermediation*, the *Journal of International Economics*, and the *Review of Finance* and it is regularly presented in leading international conferences. She earned a Master's degree and a doctorate in Economics from Boston College.

Steven Ongena is a professor in banking at the *University of Zurich* and the *Swiss Finance Institute*. He is also a research professor at *Bangor University* and a research fellow of *CEPR*. He has published more than 40 papers in refereed academic journals, including in the *American Economic Review, Econometrica, Journal of Finance, Journal of Financial Economics, Journal of International Economics*, and *Review of Finance*, among other journals, and he has published more than 40 papers in books and other collections. He is currently a co-editor of the *Review of Finance and of the International Review of Finance*; and he serves as an associate editor for a number of other journals. He is a director of the *European Finance Association* and of the *Financial Intermediation Research Society*. In 2009 he received a *Duisenberg Fellowship* from the *European Central Bank* and in 2012 a *Fordham-RPI-NYU Stern* Rising Star in Finance Award.

Chapter 10
A New Look at Bank-Firm Relationships and the Use of Collateral in Japan: Evidence from Teikoku Databank Data

Arito Ono, Hirofumi Uchida, Souichirou Kozuka and Makoto Hazama

Abstract Employing a unique micro dataset on the financial relationships between Japanese firms and their main banks and the use of collateral in their debt financing, this chapter provides a detailed account of the current landscape of business financing in Japan. The findings can be summarized as follows. First, main bank relationships are stable for most firms: less than 1 % of firm switch their main bank in any particular year, although more than 80 % of firms have established relationships with multiple banks. Second, main bank relationships are stronger in terms of deposit transactions than in terms of borrowing: the share of deposits with the main bank in the total amount of deposits is larger than the share of the amount borrowed from the main bank in the total amount of borrowing outstanding. Third, the most frequently pledged type of collateral is real estate property. And fourth, more than 30 % of realestate

The authors would like to thank Teikoku Databank Ltd. for providing data and Masahiro Miyatani and Takurou Kitou, who provided us with a detailed explanation of the TDB database. The authors also thank Noriyuki Yanagawa, Daisuke Tsuruta, Koji Sakai, and Takashi Hatakeda for their helpful comments. The views expressed in this chapter are those of the authors and do not necessarily reflect those of the institutions with which they are affiliated.

A. Ono (✉)
Chuo University, 742-1 Higashinakano, Hachioji-shi, Tokyo, Japan
e-mail: a-ono@tamacc.chuo-u.ac.jp

H. Uchida
Kobe University, 2-1 Rokkodai, Nada, Kobe, Japan
e-mail: uchida@b.kobe-u.ac.jp

S. Kozuka
Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo, Japan
e-mail: souichirou.kozuka@gakushuin.ac.jp

M. Hazama
Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, Japan
e-mail: ed104005@g.hit-u.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_10

properties are used as collateral for multiple secured loans, suggesting that the use of junior liens is quite common in Japan.

Keywords Bank-firm relationships · Main bank · Collateral

10.1 Introduction

This chapter provides a comprehensive overview of the current status of bank-firm relationships in Japan, with special emphasis on the role of main banks and the use of collateral in debt financing. This is done based on a unique and extensive dataset of Japanese firms constructed by the authors in cooperation with Teikoku Databank, Ltd. (TDB), one of the largest credit information providers in Japan. The dataset comprises nearly 400,000 Japanese firms, containing information on their basic characteristics and financial statements, transactions with suppliers and customers, and financial transactions with financial institutions (FIs).[1]

The novel features of the dataset are twofold. First, for each firm in the dataset, it is possible to identify the FIs that the firm transacts with. These FIs include both deposit-taking FIs (referred to simply as "banks" hereafter) as well as non-banks. It is also possible to identify a firm's "main bank," as subjectively recognized by the firm. Furthermore, for each firm–FI pair, it is possible to determine the amount of borrowing and time deposits outstanding. The second and more important novel feature of the dataset is that it contains detailed information on firms' use of collateral in their debt financing. For instance, it is possible to identify the types of collateral that a firm has pledged to its lenders. In addition, for each real estate property owned by a firm or its CEO and pledged as collateral, it is possible to identify the corresponding creditor and loan amount.

The rest of the chapter is organized as follows. The next section briefly describes the original TDB data and the sample selection process employed here. Section 10.3 provides descriptive statistics and univariate analyses of bank-firm relationships, while Section 10.4 describes the use of collateral. Section 10.5 concludes.

10.2 Data and Sample Characteristics

The original data are from TDB credit reports on Japanese firms for 2007–2010. Through their regular credit research, TDB collects various kinds of information related to the creditworthiness of firms, including their financial statements. Collection is conducted either in response to requests from customers (for a fee) or is

[1] A detailed description of the basic characteristics of firms in the TDB database and their inter-firm relationships, with special emphasis on the use of trade credit, can be found in Uchida et al. (2015).

unsolicited. Unsolicited data collection is conducted on leading, well-established, and/or large firms in Japan.

The sample for the present study was selected from the TDB database based on several criteria. First, the original data were available for 2007–2010, so there are firms for which data are available for multiple points in time. However, because the time span from 2007 to 2010 is too short for any meaningful time-series analysis (and data for all years are available only for some firms), the analysis focuses on a snapshot of bank-firm relationships in Japan using only the most recent data of such firms. In this sense, the dataset is cross-sectional.

Second, firms belonging to not-for-profit industries, such as religious and educational organizations, were excluded.[2] Moreover, in terms of the legal form of the firms, the sample was restricted to joint stock companies (*kabushiki-gaisha* in Japanese), closely-held limited liability companies (*yugen-gaisha*), limited partnership companies (*godo-gaisha*), limited and unlimited liability partnership companies (*goshi-/gomei-gaisha*), limited liability partnerships (*yugen-sekinin-jigyo-kumiai*), medical associations (incorporated, *iryo-hojin*), cooperative partnerships (unincorporated, *kyodo-* or *kyogyo-kumiai*), and sole proprietorships (*kojin*).

Following this sample selection process, the sample used for the analysis contains nearly 400,000 firms. Note, however, that the number of observations (denoted as "N" hereafter) for each variable differs depending on the number of missing values. In particular, the number of observations for variables based on financial statements data is significantly smaller than that for variables based on data from the TDB credit reports, since many small firms do not publish financial statements. In addition, variables from the financial statements often take implausible values. Therefore, observations were excluded if the value of a variable fell into the top or bottom 0.1 % tail.

To understand the nature of bank-firm relationships and the use of collateral in Japan, descriptive statistics of a number of variables will be examined. In addition, the results of univariate analyses, i.e., descriptive statistics when sample firms are split based on the values of various variables, are reported. It should be noted that in the univariate analyses, observations for the two variables being considered may be for different points in time. Firms were included in the analysis only if the observations for the two variables were no more than 36 months apart.

To save space, only the main findings of the analyses will be reported below, and some results will be reported without showing the associated tables or figures. The interested reader is referred to Ono et al. (2010) for full details of the analysis.

Before turning to the main findings, Table 10.1 provides the basic characteristics of the firms in the dataset.[3] More than 90 % of the firms have 100 or fewer employees,

---

[2] Note that since the aim of the analysis is to examine bank-firm relationships, financial firms such as banks, securities firms, insurance companies, and finance companies were excluded in the compilation of the dataset.

[3] See Uchida et al. (2015) for further details.

Table 10.1 Basic characteristics of firms

	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
Number of employees	372,947	49.447	426.159	0	0	10	608	140,846
Firm age	393,695	27.626	17.662	0	1	26	68	130
TDB credit score	367,224	47.524	7.538	1	30	48	66	88
Capital/total assets	247,337	0.211	0.506	− 9.333	− 1.608	0.217	0.927	0.990
Operating profit/ operating revenue	182,866	0.000	0.178	− 6.168	− 0.414	0.011	0.232	0.542
Interest bearing debt/(operating revenue/12)	165,196	5.738	13.360	0.000	0.000	3.054	52.010	374.022

so the majority of the firms are small and medium-sized enterprises (SMEs).[4] Further, it is worth noting that more than 60 % of the firms are owner-managed.[5]

10.3 Relationships Between Firms and Financial Institutions

This section presents the findings on firms' relationships with their FIs, with special emphasis on firms' relationship with their main bank(s).[6]

10.3.1 Number of Financial Institutions

Table 10.2 presents summary statistics for the number of FIs with which firms engage in lending and/or deposit transactions. There are many types of FIs: city banks, regional banks, second-tier regional banks, Shinkin banks (*shinyo kinko*), credit cooperatives (*shinyo kumiai*), long-term credit banks (LTCBs) and trust banks, other banks, government affiliated FIs, and non-banks (e.g., non-deposit taking private FIs such as finance companies, factoring companies, and leasing companies).[7] The table shows that the mean of the number of FIs with which firms transact is 3.11, while the

[4] In Japan, SMEs are officially defined in Article 2, Paragraph 1 of the Small and Medium-sized Enterprise Basic Act based on the amount of capital (up to 300 million yen) or the number of regular employees (up to 300). Lower limits apply to enterprises in the wholesale, retail, and services industries.

[5] Whether a firm is owner-managed is determined by checking the surname of the owner of the majority of the capital stock of the firm and the surname of the CEO or firm representative. If the two are the same, it is assumed that the firm is owner-managed.

[6] There already exist many studies on main bank relationships in Japan, especially on large, listed firms (see, e.g., Aoki et al. 1994; Uchida and Udell 2010).

[7] A detailed discussion of the different bank types in Japan can be found in Uchida and Udell (2010).

Table 10.2 Number of financial institutions

Number of financial institutions	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	373,155	3.114	1.971	1	1	3	10	48
Distribution	1	2	3	4	5	6	7	8 or more
Freq.	68,325	98,543	86,015	53,741	29,580	15,987	8,833	12,131
(% share)	(18.31)	(26.41)	(23.05)	(14.40)	(7.93)	(4.28)	(2.37)	(3.25)

median is 3. Looking at the distribution of the number of transaction relationships, only 18 % of firms in the sample transact with only one FI, and most firms transact with two to four FIs. As noted in Table 4.4 of Degryse et al. (2009), which provides a summary of empirical studies on the number of bank relationships in countries around the world, multiple bank relationships are common not only in Japan, but also in other countries.

Looking at the number of FIs that firms transact with, this differs depending on firm characteristics (no table provided to conserve space). First, small firms have fewer relationships than large firms. For example, the average (mean) number of FIs for firms that have no employees is 2.1, whereas that for firms with more than 1000 employees is 5.8. This finding is consistent with findings for other countries (Degryse et al. 2009). Second, firms with a low TDB credit score, i.e., riskier firms, tend to have fewer transaction relationships with FIs.[8] Third, firms that transact with government FIs or that rely on government-funded credit guarantee programs[9] tend to transact with a larger number of FIs. The average number of FIs for firms whose main bank is a government FI is 4.2 (compared to 3.2 for firms whose main bank is not a government FI), and that for firms using a government-funded credit guarantee program is 3.5 (compared to 2.8 for non-users). In sum, firms that transact with a large number of FIs tend to be either creditworthy firms with a high credit score or firms that require government support, such as credit guarantees, to obtain loans.[10]

[8] The TDB score is a metric that evaluates the creditworthiness of a firm. The score ranges from 1 to 100 points, and a higher score indicates higher creditworthiness. See Uchida et al. (2015) for more on TDB scores.

[9] To mitigate the financing difficulties of SMEs, the Japanese government employs credit guarantee programs that ensure the repayment of defaulted loans. See Uesugi et al. (2010) and Ono et al. (2013) for details on credit guarantee programs in Japan.

[10] However, (and as will be seen below), firms whose main bank is a government FI are not necessarily riskier.

10.3.2 Main Banks

10.3.2.1 Composition

The TDB database contains information on the main bank of firms in the sample used here. Main bank relationships are multifaceted and involve the provision of various financial services. The definition of a main bank in the TDB database is somewhat subjective because it is based on identification by the firms themselves. Formally, TDB defines a main bank as the bank that provides the largest amount of loans for working capital.[11] This means that in some cases the main bank as identified by the firm itself and the one based on the TDB definition are not identical. In these cases, the former takes precedence and is identified as the main bank.[12]

Table 10.3 shows the composition of main banks of firms in the sample used here. Regarding the number of main banks, 90.1 % of all sample firms ($N = 310{,}097$) list just one main bank, while 8.6 % list two (not shown in the table). Looking at firms that have a single main bank, 36.7 % of those main banks are a regional bank (Table 10.3, left column). The percentage shares of city banks (27.4 %), Shinkin banks (21.0 %), and second-tier regional banks (10.6 %) are also sizable. However, when looking at firms that have more than one main bank (Table 10.3, center column), city banks make up the largest share (42.7 %), since such firms are likely to be relatively large and are thus likely to transact with larger banks.

Three further aspects are worth mentioning. First, large firms (in terms of the number of employees) tend to have city banks, LTCBs, and trust banks as their main bank, whereas Shinkin banks and credit cooperatives are more likely to be the main bank for smaller firms. Second, firms with a higher TDB credit score, i.e., firms with lower credit risk, are more likely to have a larger bank as their main bank. Third, firms that have a government FI as their main bank tend to be larger and have higher credit scores. For instance, the mean number of employees (71.3) and the mean TDB credit scores (49.6) for such firms are comparable to those for firms with a city bank as their main bank (82.1 employees and a TDB credit score of 49.7). This suggests that main bank relationships with government FIs are substitutes for rather than complements to relationships with large private FIs.

---

[11] More precisely, a main bank is a bank with which a firm has a deposit account and obtains loans (short-term and/or long-term loans). When a firm has a deposit relationship with one bank and a lending relationship with another, TDB identifies the latter as the main bank, even though the firm is allowed to list multiple banks as its main banks. Similarly, when a firm obtains short-term loans from one bank and long-term loans from another, then the former is labeled as the main bank.

[12] In the empirical literature, the main bank is typically defined as the private FI with which a firm has the largest amount of outstanding loans (Kawai et al. 1996; Sheard 1989); however, other definitions are also used (see Uchida and Udell 2010).

Table 10.3 Composition of main banks

	Composition for firms with a single main bank		Composition for firms with multiple main banks		Composition of all main banks	
	Freq.	(% share)	Freq.	(% share)	Freq.	(% share)
City banks	76,666	(27.44)	28,095	(42.65)	104,761	(30.34)
Regional banks	102,647	(36.73)	17,121	(25.99)	119,768	(34.68)
Second-tier regional banks	29,685	(10.62)	5,603	(8.51)	35,288	(10.22)
Shinkin banks	58,545	(20.95)	9,855	(14.96)	68,400	(19.81)
Credit cooperatives	6,985	(2.50)	1,351	(2.05)	8,336	(2.41)
Long-term credit & trust banks	420	(0.15)	778	(1.18)	1,198	(0.35)
Government FIs	2,786	(1.00)	2,586	(3.93)	5,372	(1.56)
Other banks	1,690	(0.60)	478	(0.73)	2,168	(0.63)
Non-banks	12	(0.00)	10	(0.02)	22	(0.01)
Total number of main banks	279,436	(100.00)	65,877	(100.00)	345,313	(100.00)
Total number of firms	279,436		30,661		310,097	

10.3.2.2 Switching Main Banks

Previous studies on main bank relationships in Japan found that firms rarely switch their main bank. For instance, Kano (2007) found that only 15 % of Japanese SMEs switched their main bank during 1980–1990 and 1990–2000 (i.e., 1.5 % per annum).

The TDB database identifies whether a firm has switched its main bank within the past 2 years. The results suggest that 4,610 of the 373,599 firms (1.2 %) in the sample switched main banks. Thus, the annual rate, 0.6 %, is smaller than, but comparable to, that in Kano (2007).

Regarding the characteristics of firms that switched main bank, there is a non-monotonic relationship between firm size (represented by the number of employees) or the riskiness of the firm (represented by TDB credit scores) on the one hand and the likelihood of switching on the other: firms with 10–50 employees and those with TDB scores of 40–50 are more likely to switch.

10.3.3 *Transactions Between Firms and Financial Institutions: Borrowing*

10.3.3.1 Total Amount of Borrowing

The TDB database compiles information about firms' financial transactions in terms of borrowing and time deposits.[13] With respect to the amount of borrowing, the mean and the median amounts are 2.97 billion yen and 79 million yen, respectively ($N = 323{,}847$). If firms that fall into both 0.1 % tails of the sample distribution are excluded, the mean becomes much smaller: 596 million yen. Moreover, 27,739 firms (8.6 % of the sample firms) have no borrowing outstanding.

Borrowing comes from three sources (Table 10.4). The main source (lenders) is FIs. The mean and median amounts of borrowing from FIs are 2.46 billion yen and 72 million yen, respectively (N = 306,295). Among firms that reported the amount of borrowing from FIs (including firms that reported zero for borrowing), 29,551 (9.6 %) have no borrowing outstanding. Second, some firms obtain loans from "insiders," such as CEOs, executives, and affiliate companies, including both subsidiaries and parents. Among the 215,206 firms that reported the amount of borrowing from insiders, more than half (54.3 %) have positive borrowing outstanding from insiders, while the rest reported zero for borrowing from insiders. Looking at the ratio of borrowing from insiders to total borrowing, 17.5 % of the firms are almost exclusively dependent on insider finance (i.e., borrowing from insiders accounts for 95–100 % of total borrowing). The mean and median amounts of outstanding borrowing from insiders are 489 million yen and 2 million yen, respectively, and are far smaller than the mean and median of outstanding borrowing from FIs. Third, 134,124 firms disclosed the amount of corporate bonds outstanding, including both public bonds and private placements. The ratio of firms that use corporate bonds is 17.0 %, and the average amount of bonds outstanding is 545 million yen. The mean amount of borrowing from insiders and corporate bonds is considerably smaller if firms that fall into both 0.1 % tails of the sample distribution are excluded.

Turning to the univariate analysis, this shows that the average amount of borrowing is proportional to firms' size, their creditworthiness, and the size of their main bank. One interesting finding is the unique characteristics of firms that do not have any borrowing outstanding. Figure 10.1 shows the share of firms with no borrowing outstanding for each firm size (number of employees: left) and level of creditworthiness (TDB credit score: right). The left panel shows that among firms with no employees ($N = 16{,}672$), 13.1 % have no borrowing outstanding. Moreover, the figure exhibits a U-shape, implying that it is primarily the smallest and largest firms that have no borrowing outstanding. For small firms, the fact that they have no borrowing outstanding may be due to difficulties in obtaining external funds, either

[13] The TDB database also contains information regarding discounted bills (*tegata-waribiki*), a traditional method for short-term financing in Japan. See Ono et al. (2010) for more on discounted bills.

Table 10.4 Amount of borrowing (from FIs, insiders, and through bonds) and time deposits

(Million yen, except N)

	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
Total borrowing	323,847	2,970	273,000	0	0	79	8,800	79,800,000
(w/o 0.1 % tails)	(323,162)	(596)	(4,050)	(0)	(0)	(79)	(8,000)	(180,000)
Borrowing from FIs	306,295	2,460	245,000	0	0	72	7,490	78,700,000
(w/o 0.1 % tails)	(305,669)	(512)	(3,240)	(0)	(0)	(72)	(6,870)	(140,000)
Borrowing from insiders	215,206	489	35,600	0	0	2	2,100	8,600,000
(w/o 0.1 % tails)	(214,740)	(100)	(806)	(0)	(0)	(2)	(1,930)	(28,800)
Corporate bonds	134,124	545	69,400	0	0	0	1,230	25,100,000
(w/o 0.1 % tails)	(133,850)	(100)	(1,390)	(0)	(0)	(0)	(1,100)	(71,000)
Time deposits	51,323	627	63,600	0	0	25	1,800	14,000,000
(w/o 0.1 % tails)	(51,220)	(134)	(683)	(0)	(0)	(25)	(1,640)	(25,100)

Note: "FIs" stands for financial institutions. "Insiders" include CEOs, executives, and affiliates.

because their loan applications had been declined or they had been discouraged from applying for a loans to begin with. For large firms, the result is consistent with the casual observation that large firms can rely on internal funds such as operational cash flow rather than external debt. Regarding firms' creditworthiness, the figure (right) shows that firms with a higher credit score are less likely to borrow at all. Looking at the share of firms with no borrowing outstanding by firm age (not reported), this is slightly higher for younger firms.

As for borrowing from insiders (CEOs, executives, and affiliates), it is smaller, riskier, and younger firms that tend to depend more heavily on loans from insiders.

10.3.3.2 Borrowing from Main Banks

Where possible, the TDB database lists the identities of the FIs that a firm transacts with, including both its main bank(s) and non-main banks, together with the amount of borrowing outstanding from each FI. Of the firms in the database, only 92,489 firms report the amount of borrowing outstanding from their main bank(s) (including firms that report zero borrowing), which is less than one-third of the number of firms for which the amount of total borrowing is available ($N = 323{,}847$).

The mean and median values of the amount of borrowing outstanding from the main bank are 3.53 billion yen and 50 million yen, respectively (Table 10.5). It is important to note that 31.5 % of these firms do not have any outstanding borrowing from their main bank. For reference, the share of firms that do not borrow at all from any FI is only 0.3 %. Thus, there are many firms that only have lending relationships with non-main banks. This sounds peculiar, given that a main bank is defined as the bank that provides loans for working capital to the firm. Although we cannot pin down the reason for this inconsistency, there are several possible explanations. First, even if a firm does not borrow at all from its main bank, a firm might set up a credit

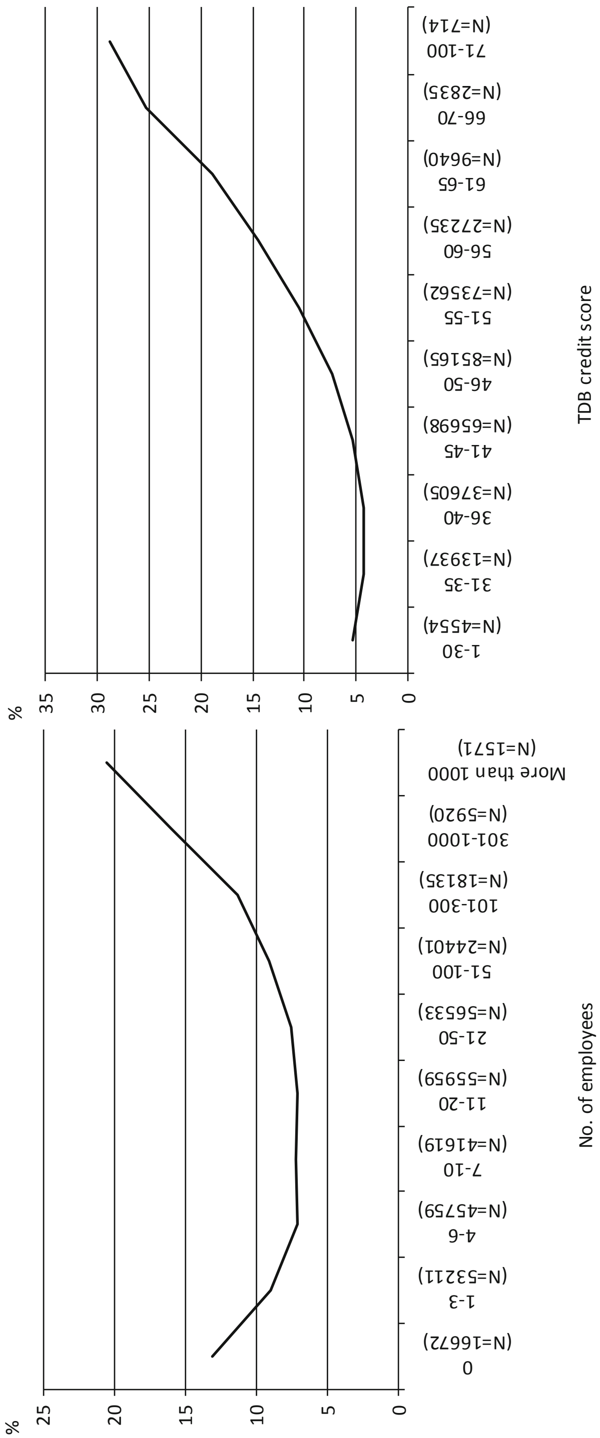


Fig. 10.1 Percentages of firms with no borrowing outstanding: Univariate analysis

Table 10.5 Transactions with main banks: Amount and concentration ratios

	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
Amount (in million yen)								
Borrowing	92,489	3,530	424,000	0	0	50	3,810	81,900,000
(w/o 0.1 % tails)	(92,304)	(274)	(1,450)	(0)	(0)	(50)	(3,500)	(63,000)
Time deposits	36,927	188	15,000	0	0	15	630	2,730,000
(w/o 0.1 % tails)	(36,854)	(55)	(143)	(0)	(0)	(15)	(600)	(4,000)
Concentration ratio (%)								
Borrowing	91,733	0.450	0.396	0.000	0.000	0.429	1.000	1.000
(w/o 0.1 % tails)	(91,610)	(0.450)	(0.396)	(0.000)	(0.000)	(0.429)	(1.000)	(1.000)
Time deposits	36,520	0.658	0.384	0.000	0.000	0.808	1.000	1.000
(w/o 0.1 % tails)	(36,447)	(0.659)	(0.384)	(0.000)	(0.000)	(0.808)	(1.000)	(1.000)

line (including an overdraft arrangement) with a main bank. Second, firms might not define their main bank solely based on lending relationships. It might be the case that firms take into account different financial transactions, especially deposit relations, or that they define their main bank as one that they think most reliable in times of difficulty.

Computing the concentration ratio of main banks, defined as the amount of borrowing outstanding from the main bank divided by the amount of borrowing from all FIs (Table 10.5) shows that the mean and the median of the concentration ratio are 45.0 and 42.9 %, respectively ($N = 91{,}733$). When confined to short-term loans, the mean concentration ratio is 19.6 %, whereas the ratio for long-term loans is 26.5 %.

Next, looking at the concentration ratio of borrowing from the main bank in terms of different firm characteristics shows the following. First, smaller firms (in terms of the number of employees) tend to depend more heavily on their main bank than larger firms, as indicated by the fact that 22 % of firms with less than 4 employees have no borrowing outstanding from their main bank compared to 45.1 % of firms with more than 300 employees. Second, the concentration ratio is higher for firms with a lower TDB credit score. This implies that creditworthy firms depend less on their main bank. This finding is consistent with the finding above that creditworthy firms are more likely not to borrow at all (Fig. 10.1).

10.3.4 Transactions Between Firms and Financial Institutions: Deposits

10.3.4.1 Total Amount of Time Deposits

For a small number of firms, the TDB database contains information regarding the amount of time deposits. In contrast to the number of firms for which the total amount of borrowing is available, the number of firms for which information on time deposits is available is very small, yielding a sample size of only 51,323 firms (Table 10.4). For these firms, the mean and median amounts of time deposits outstanding are 627 million yen and 25 million yen, respectively.

The share of firms that have no time deposits is 13.8 %. Moreover, the share of firms with no time deposits is larger among the smallest and largest firms, yielding a U-curved relationship between firm size and the likelihood of having time deposits. On the other hand, firms with a lower credit score are more likely to have no time deposits. The latter finding suggests that firms with a lower score are more likely to be liquidity constrained and therefore do not have time deposits outstanding.

10.3.4.2 Transactions with Main Banks

In addition to the summary statistics for the amount of borrowing, Table 10.5 also presents those for the amount of time deposits outstanding that a firm has vis-à-vis its main bank. The number of observations for main bank transactions is again small, approximately 37,000. Table 10.5 also compares the main bank concentration ratio in terms of borrowing and time deposits. As noted above, the primary definition of a main bank in the TDB credit reports is based on lending relationships. However, the average concentration ratio of transactions with the main bank is larger for time deposits (65.8 %) than for borrowing (45.0 %). These ratios seem to be inconsistent with TDB's formal definition of a main bank.

To investigate this inconsistency further, Fig. 10.2 compares the distribution of the main bank concentration ratios in terms of borrowing and time deposits. The figure shows that both distributions are extremely polarized, with the largest shares of firms either having a very low or a very high concentration, although the degree differs for borrowing and time deposits. Specifically, both distributions are asymmetric, and in the case of time deposits the share of firms with no such deposits outstanding with their main bank (16.5 %) is much smaller than that of completely main bank-dependent firms (i.e., firms with a concentration ratio equal to 95–100 %, which make up 43.8 %). In contrast, the share of firms with no borrowing outstanding from their main bank (30.4 %) is higher than that of fully main bank-dependent firms (14.7 %). As a result, the average main bank concentration ratio for time deposits is higher than that for borrowing.

To summarize, the results here indicate that firms' reliance on their main bank—as measured in terms of the main bank's share in a firm's total borrowing or deposits—is more pronounced for deposit transactions than for lending transactions. While the literature on main bank relationships and relationship banking for SMEs largely focuses on lending relationships, this finding suggests that future research should also focus on deposit transactions.[14]

[14] Mester et al. (2007) provide empirical evidence on the importance of deposit relationships (transaction accounts) in monitoring borrowers.

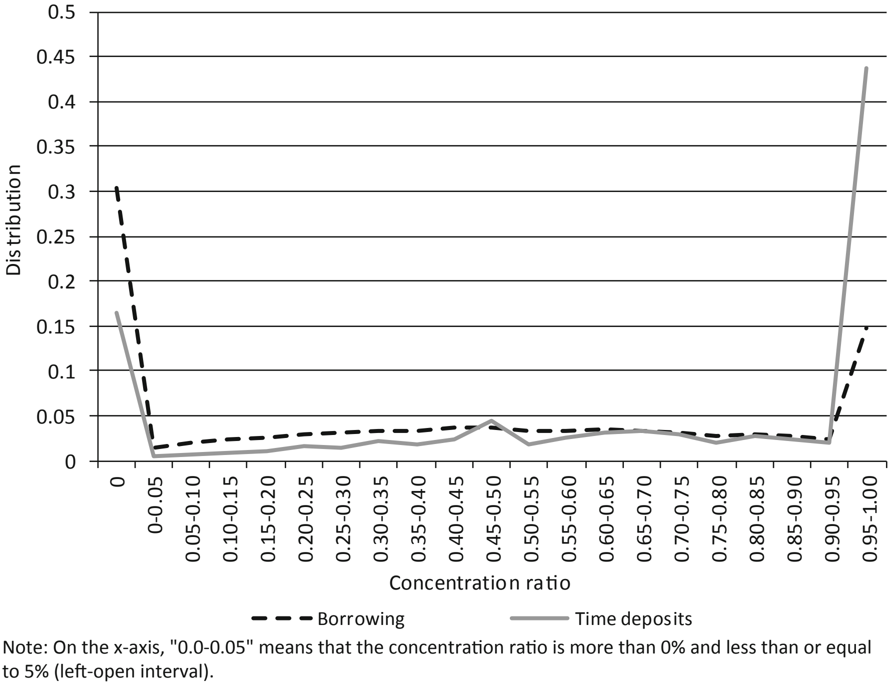


Fig. 10.2 Distribution of main bank concentration ratios

10.4 Collateral

This section reports the findings on the use of collateral.

10.4.1 Data

Data on the use of collateral comes from three sources of information in the TDB database. First, in addition to the information on the bank-firm relationships reported in Section 10.3, TDB credit research also shows the various types of collateral pledged by firms for their borrowing (Source 1). Second, TDB also collects information on whether a firm registers its account receivables and/or inventories as collateral with the public registry in Japan (Source 2). Third, and most importantly, the TDB database contains very detailed information on the properties of a firm that are registered with the public real estate registry (Source 3).

Although the TDB database contains rich data on collateral compared to other databases in Japan and elsewhere, several caveats should be noted. First, with the exception of the data from the public real estate registry (Source 3), the identity of the secured lender is not available. For instance, when a firm pledges a piece of real estate to bank A and pledges some securities (e.g., stock certificates) to bank B, the TDB database only shows that the firm uses real estate and securities as collateral and does not provide information regarding to which lender the collateral was pledged. Second, data from the public real estate registry (Source 3) suffer from sample selection bias problems. The reason is that, in its credit research, TDB puts much more emphasis on real estate assets in the case of SMEs than larger firms. For SMEs, TDB always obtains registered information on CEOs' residential real estate and the land and buildings that constitute a firm's headquarters, because these properties are most likely to be pledged as collateral in business loans to SMEs. However, information regarding the CEO's or the firm's other real estate is obtained only on request. Furthermore, for large firms (with paid-in capital of more than 100 million yen and with more than 100 employees), the TDB does not obtain such information unless requested by customers.[15]

10.4.2 Types of Collateral

10.4.2.1 Composition

The TDB database contains information on whether a firm uses different types of assets as collateral or guarantees (from Source 1): real estate (owned by a firm, the CEO, or others), securities, deposits, credit guarantees by private firms (such as parent companies or guarantee companies), government credit guarantees, and other assets (such as equipment, account receivables, and inventories). Note that the database does not provide information on personal guarantees provided by individuals (e.g., CEOs and executives), which are frequently used in Japan.[16]

Figure 10.3 shows the percentage of firms that use each type of collateral or guarantees ($N = 373{,}599$). Of the total sample firms, 23.5 % did not pledge any collateral or guarantees, and 51.9 % pledged real estate collateral. Among real estate, the most widely used is properties owned by the firm (internal assets). One-quarter of firms pledged the CEO's properties (external assets) as collateral. Following real estate, deposits are the second most frequently used assets as collateral.[17] The prevalence

[15] The percentage of firms with paid-in capital of more than 100 million yen in the real estate registry data is approximately 3 %, whereas it is 8 % in the full sample used here. Similarly, the percentage of firms with more than 100 employees in the real estate registry data is approximately 6 %, while it is 8 % for the full sample.

[16] Ono and Uesugi (2009) showed that 66.7 % of Japanese SMEs use personal guarantees.

[17] In Japan, most firms open a checking account with their main bank, because promissory notes or checks are accepted and discounted only when drawn against a checking account with a bank and produced by using the uniform format promulgated by the bankers' association. Once such a checking account is opened, a large part of the firm's cash flow goes through the account. Therefore, taking deposits in the checking account as collateral could be seen as equivalent to collateralizing the cash flow of the firm.

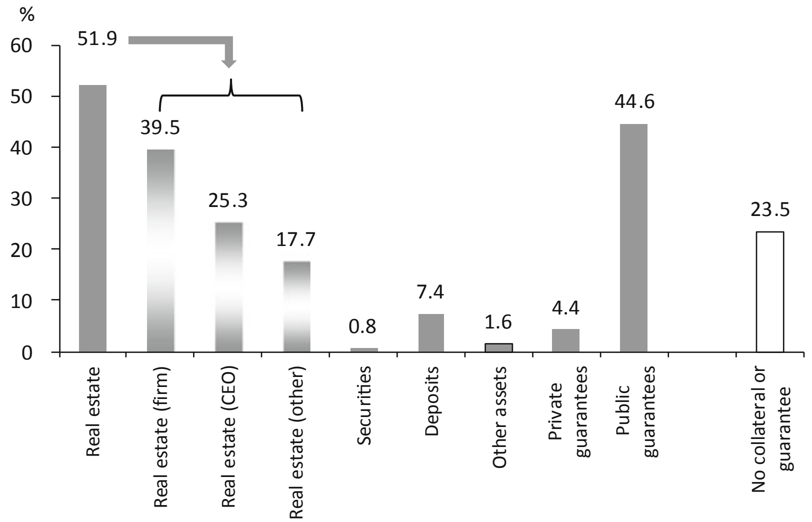


Fig. 10.3 Percentage of firms that pledged collateral and guarantees: by type of security pledged

of public credit guarantees is also noteworthy, given that 44.6 % of all firms obtained guarantees backed, at least partially, by the government to obtain external debts.

For a limited number of firms, information is also available on whether they register their account receivables and inventories as collateral (obtained from Source 2). Because it is not mandatory for TDB researchers to obtain such information, the number of observations (firms) is limited: 27,310 for account receivables and 27,280 for inventories.

The percentages of firms that register account receivables and inventories as collateral are 3.3 and 0.9 %, respectively. However, assuming that firms for which the data entry is "N.A." have not pledged these types of collateral, the percentages of firms pledging account receivables and inventories as collateral decrease to 0.2 and 0.1 %. The true percentages probably lie somewhere between these two extremes.[18] Following the collapse of the real estate bubble in the early 1990s and the financial crisis in Japan in the late 1990s, it was widely expected that asset-based lending using account receivables and inventories as collateral would help to increase business loans to firms that do not have sufficient real estate to pledge as collateral.[19] However,

[18] In Fig. 12.3, account receivables and inventories are included in the item "other assets."

[19] Following the tradition of the Civil Law system, Japanese law does not allow non-possessory security interests in movables. The Civil Code stipulates that movables cannot be the subject of a mortgage, but can only be the subject of a pledge, which requires that the debtor (pledgor) gives possession of the movable to the creditor (pledgee). However, in order to enable the debtor to utilize their property as a security for their debt and still use the same property for the business, the practice of "transfer by way of security" has developed. This is the legal practice of transferring the ownership of the movable to the creditor based on the agreement that the transfer is only for the sake of securing the credit and that ownership will revert to the debtor when the latter has

as indicated by the small percentages of firms that register account receivables and inventories as collateral, asset-based lending at present is still underdeveloped.

Comparing the characteristics of firms that register account receivables and inventories and those that do not, it appears that larger firms tend to pledge these more frequently as collateral. For instance, the percentage of firms with more than 300 employees that register account receivables is 2.5 % (in the case when we consider "N.A." as no collateral), whereas it is 0.2 % for firms with 300 or fewer employees.

10.4.2.2 Univariate Analysis

Figure 10.4 presents the univariate analysis for the percentage of firms that use real estate collateral, the most frequently used collateral in Japan. Interestingly, there is no monotonic relationship between the use of real estate collateral and firm size or credit risk. With respect to firm size, the share of firms using real estate as collateral increases with the number of employees for firms with up to 50 employees. It then remains more or less unchanged as firm size increases between 50 and 100 employees, and then decreases again for larger firms. The figure for TDB credit scores similarly shows an inverse U-shape, and reaches its maximum for firms with a credit score between 50 and 55. One possible interpretation of these findings is that smaller and riskier firms do not have real estate to pledge, whereas larger and more creditworthy firms do not need any collateral to obtain loans. Comparing the use of real estate collateral by firms' industry, manufacturing (66.3 %) and real estate (61.5 %) firms are more likely to pledge real estate, whereas firms in the service sector (34.9 %) are less likely to do so. Moreover, comparing the use of collateral by bank type, the rate of collateral usage is approximately 40 % for firms whose main bank is a large bank such as a city bank, an LTCB, or a trust bank, while it is more than 60 % for firms whose main bank is a regional bank, a second-tier regional bank, a Shinkin bank, or a credit cooperative. It is also noteworthy that for firms whose main bank is a government FI, nearly 80 % pledged real estate as collateral.

repaid the debt. In 1998, a registration (filing) system for account receivables was introduced with the expectation that registration enhances the effectiveness of transfers by way of security. Under the new registration system, the security interest is "perfected," i.e. becomes valid vis-à-vis the third party (another creditor). In 2004, the registration system was extended to movables (inventories and equipment). Notwithstanding these reforms, the system has several limitations. In particular, registration is not the exclusive means for the perfection of the interest in a transfer by way of security; as another means of perfecting the creditor's right remains taking possession of the property. Therefore, registration does not preclude the possibility that another creditor claims an interest that has priority over the registered creditor based on the fact that the unregistered creditor had taken possession of the property (typically through the "fictitious possession" by the debtor's declaring that the debtor possesses the movable for the sake of the unregistered creditor) prior to the registration. As a result, a creditor cannot be sure whether they can rely fully on a registration to secure the loan (Kozuka and Fujisawa 2009). The low level of account receivables and inventories registered as collateral may partly be the result of such shortcomings of the registration system.

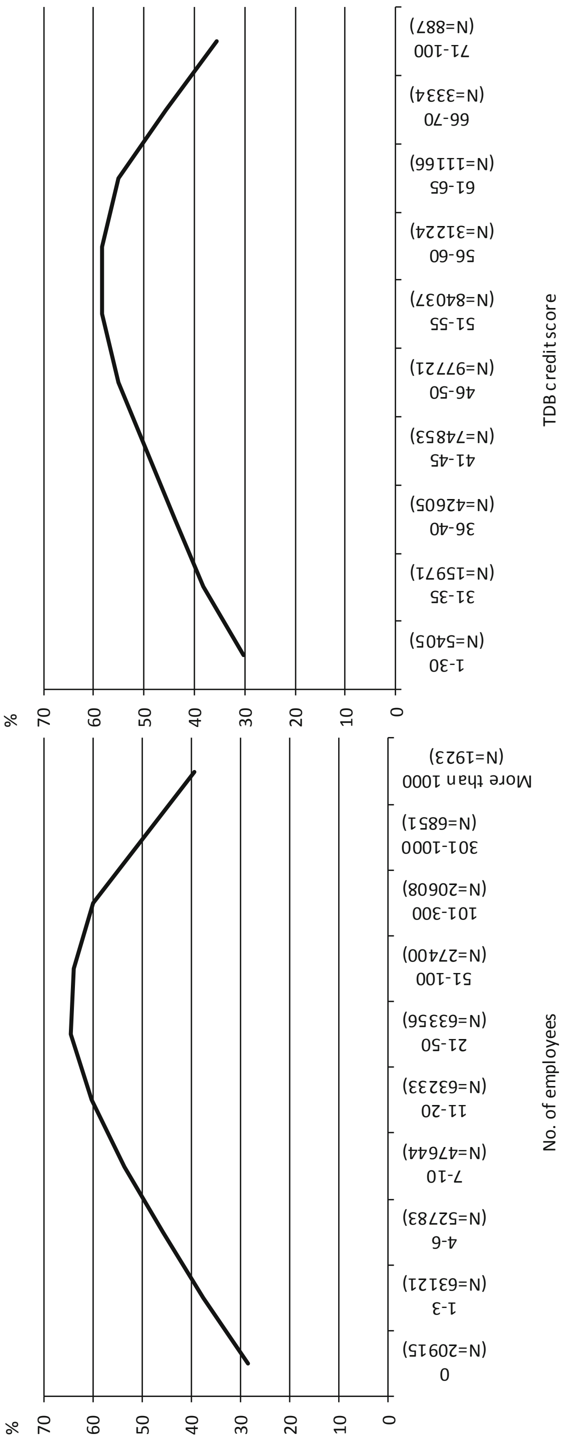


Fig. 10.4 Percentage of firms that pledged real estate: Univariate analysis

10.4.3 Real Estate Collateral

This section provides a detailed description of real estate collateral using information obtained from the public real estate registry (Source 3).

10.4.3.1 Number of Properties, Properties Pledged, and Secured Loans

Among the 310,058 firms shown in Table 10.6 (upper panel), the mean and median values for the number of properties owned by a firm are 8.8 and 6, respectively. Because land and buildings are usually counted separately in the real estate registry, a value of 2 effectively means one piece of land plus one building on it.[20] Looking at the distribution of the number of properties owned, the most frequent observation (mode) is 2 (17.3 %), followed by 4 (9.7 %). This suggests that less than 20 % of the firms have only one piece of land with one building, probably the headquarters or the CEO's residential real estate, and approximately 10 % have two pieces of land/building.[21]

Table 10.6 also shows that 257,829 firms have at least one property pledged as collateral, with a mean of 7.4 and a median of 5 (Table 10.6, middle panel). Taking the ratio of the number of properties pledged to the number of properties owned, the mean is 78.3 % and the median 86.7 %, implying that on average, 78–87 % of the properties owned by firms serve as collateral (Table 10.6, lower panel). For 39.6 % of all firms, the share of properties they have pledged as collateral is 100 %; that is, these firms have pledge all the properties they own to lenders. Note that in the case of residential real estate, this figure might include both security interests for business loans and residential loans.

Regarding the number of secured loans per firm, the mean is 3.3 and the median is 2 ($N = 258{,}012$, Table 10.7 upper panel). Of the total sample, 31.0 % of firms have only one secured loan, and 22.8 % have two (Table 10.7, upper panel).

The data also show that the average number of properties pledged for one secured loan is 5.5, while the median is 3, although the mode is 2 (Table 10.7, middle panel, $N = 840{,}898$). Next, calculating the number of secured loans per property, i.e., the number of secured loans to which one property is pledged, (Table 10.7, lower panel, $N = 2{,}743{,}604$) shows that this is 0 for 30.4 % of the total properties, meaning they are not used as collateral. The share of properties used for one secured loan is 36.6 %, while the share of properties used for multiple secured loans is 33 %. Given that for properties secured by multiple loans there needs to be an order of priority among these loans, the fact that more than 30 % of properties are used for multiple secured loans suggests that the use of junior lien loans is common practice in Japan.

[20] In cases where land and building are not separable (as in the case of an apartment), this is counted as one piece of real estate.

[21] Note, however, that the share of firms with four or fewer properties is probably overestimated, since it is possible that TDB does not investigate whether a firm has other properties.

Table 10.6 Number of properties owned and pledged as collateral

No. of properties	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	310,058	8.849	9.167	1	1	6	44	257
Distribution	1	2	3	4	5	6	7	8
Freq. (% share)	5,053 (1.63)	53,572 (17.28)	26,118 (8.42)	30,028 (9.68)	24,753 (7.98)	22,166 (7.15)	19,395 (6.26)	17,595 (5.67)
Distribution	9	10	11–20	21–30	More than 30			
Freq. (% share)	14,488 (4.67)	12,511 (4.04)	59,990 (19.35)	15,262 (4.92)	9,127 (2.94)			
No. of properties pledged as collateral	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	257,829	7.436	8.336	1	1	5	40	238
Distribution	1	2	3	4	5	6	7	8
Freq. (% share)	17,275 (6.70)	49,203 (19.08)	27,873 (10.81)	26,406 (10.24)	20,813 (8.07)	17,727 (6.88)	14,661 (5.69)	12,157 (4.72)
Distribution	9	10	11–20	21–30	More than 30			
Freq. (% share)	9,929 (3.20)	8,357 (2.70)	38,447 (12.40)	9,275 (2.99)	5,706 (1.84)			
No. of properties pledged as collateral/No. of properties	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	257,829	0.783	0.245	0.009	0.143	0.867	1.000	1.000

10.4.3.2 Secured Creditors (Lenders)

Next, for firms that use real estate properties as collateral, the number of secured creditors (i.e., lenders) is examined. Note that the number of secured creditors per firm can differ from the number of loans per firm (see Section 10.4.3.1), since a lender can have several secured loans to the same firm. Note also that lenders include not only FIs but also non-financial firms and individuals.

The mean and the median of the number of secured creditors per firm are 2.0 and 2, respectively (results not reported). Out of the 258,048 firms for which the number of secured creditors is available, 44.1 % only have one secured creditor, while 29.6 % have two creditors. Looking at the composition of secured creditors, regional banks make up the largest share (18.8 %), followed by Shinkin banks (14.0 %), city banks (10.2 %), and second-tier regional banks (6.8 %). However, these figures are smaller than the shares of these types of banks as main banks (see Table 10.3). This is because government FIs (17.5 %), non-financials (11.3 %), and private credit guarantee corporations (11.6 %)—which all are unlikely to be firms' main bank—account for sizable percentages. The reason for the high percentage of private credit guarantee

Table 10.7 Number of secured loans per firm, properties pledged as collateral per secured loan, and secured loans per property

No. of secured loans per firm	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	258,012	3.260	3.120	1	1	2	15	134
Distribution	1	2	3	4	5	6	7	8 or more
Freq. (% share)	80,025 (31.02)	58,826 (22.80)	38,611 (14.96)	24,952 (9.67)	16,725 (6.48)	11,132 (4.31)	7,773 (3.01)	19,968 (7.74)
No. of properties pledged as collateral per secured loan	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	840,898	5.499	11.162	1	1	3	37	1,076
Distribution	1	2	3	4	5	6	7	8 or more
Freq. (% share)	82,483 (9.81)	248,322 (29.53)	137,634 (16.37)	89,476 (10.64)	61,263 (7.29)	43,271 (5.15)	31,759 (3.78)	146,690 (17.44)
No. of secured loans per property	N	Mean	Std. dev.	Min.	p1	p50	p99	Max.
	2,743,604	1.485	2.061	0	0	1	9	132
Distribution	0	1	2	3	4	5	6	7 or more
Freq. (% share)	833,419 (30.38)	1,005,248 (36.64)	435,340 (15.87)	203,184 (7.41)	104,129 (3.80)	59,176 (2.16)	34,309 (1.25)	68,799 (2.51)

corporations is that secured loans in the sample include residential mortgages, which are often secured by private credit guarantee corporations affiliated with the bank that extended the loan.

Looking at the characteristics of firms that obtained secured loans from non-financial firms shows that on average such firms are smaller in size (in terms of the number of employees) and riskier (in terms of TDB's credit score). By industry, manufacturing firms are less likely to obtain secured loans from non-financial firms, whereas wholesale and retail firms (including restaurants) are more likely to do so.

10.4.3.3 Seniority Among Secured Creditors

This section examines the seniority among creditors for secured loans. To this end, a seniority index for each property pledged as collateral is constructed in the following manner. If there is only one secured creditor for a particular property, then the index takes a value of 1 if the secured creditor is the main bank of the firm, and a value of 2 otherwise. If there are multiple secured creditors for a property, then it is assumed

Table 10.8 Seniority among secured creditors

	Freq.	(% share)
1. One secured creditor, main bank	724,966	(42.75)
2. One secured creditor, non-main bank creditor	506,471	(29.86)
3. Multiple secured creditors, main bank and non-main bank creditor, both have seniority	44,004	(2.59)
4. Multiple secured creditors, main bank and non-main bank creditor, main bank has seniority	218,155	(12.86)
5. Multiple secured creditors, main bank and non-main bank creditor, non-main bank creditor has seniority	114,766	(6.77)
6. Multiple secured creditors, non-main bank creditors only	67,837	(4.00)
7. Multiple secured creditors, main banks only	19,726	(1.16)
Total	1,695,925	(100.00)

that the loan with the older registration date is senior.[22] In the case where both a main bank and a non-main bank creditor are secured, the index takes a value of 3 if they are of equal seniority, 4 if the main bank is senior, and 5 if the non-main bank creditor is senior. The index takes a value of 6 in the case where all multiple secured creditors are non-main bank creditors and 7 in the case where multiple main banks are secured.

As noted above, the sample includes residential mortgages. To remove them, observations are excluded when the debtor is the CEO of the firm and when the firm has not pledged the real estate owned by CEOs to the creditor. To avoid possible inconsistencies among different sources of information in the TDB database, observations are also excluded when an FI identified as the creditor (in Source 3) was not listed in TDB's list of FIs that the firm transacts with (in Source 1). The remaining sample consists of 200,614 firms and 1,695,925 pledged properties.

Table 10.8 shows that more than 40 % of such properties ($N = 1{,}695{,}925$) are secured by only one main bank (shown as 1 in the index). Surprisingly, nearly 30 % of the properties are only secured by non-main bank creditors, including non-financial firms (index = 2); this fact seems to be inconsistent with practitioners' view that main banks usually possess senior security interests. Meanwhile, 12.9 % of properties are pledged to both main banks and non-main bank creditors, and the main bank has the first priority (index = 4). Further, 6.8 % of properties are pledged to both main banks and non-main bank creditors and the non-main bank creditors have seniority.

[22] Although the official real estate registry provides information about seniority among secured loans, the TDB does not collect such information.

10.5 Conclusion

This chapter presented a brief overview of bank-firm relationships and the use of collateral in Japan using an unprecedented and extensive dataset. The uniqueness of this dataset made it possible to obtain several novel findings.

Regarding bank-firm relationships, it was confirmed that main bank relations in Japan are stable for most firms and that the ratio of firms that switch their main banks is less than 1 % per annum. However, more than 80 % of firms have established relationships with multiple banks. The findings also highlight the importance of deposit transactions between firms and their main banks. For instance, the average main bank concentration ratio is larger for time deposits than for borrowing. Further, it was found that not only main bank relationships, but financing from insiders is also important: more than half of the firms in the sample rely on borrowing from insiders such as CEOs, executives, and affiliates. Insider financing is especially relevant for smaller, riskier, and younger firms. In addition, it was found that government FIs do not necessarily lend to smaller or less creditworthy firms, which are likely to have difficulties in financing, and that their borrowers pledge collateral more frequently than firms that borrow from private banks.

As for the use of real estate property as collateral, the findings show that the firms most likely to pledge properties as collateral are medium-sized and medium risk firms. Focusing on borrowing from main banks, the results indicate that regional banks and cooperative financial institutions more frequently require borrowers to pledge collateral than larger banks such as city banks. Using information from the public real estate registry, it was further found that there are many cases in which firms obtain multiple secured loans based on one real estate property, implying that there is a priority order among creditors and that junior lien loans are commonly used in Japan. Moreover, non-main bank creditors often have senior security interests. Looking at the distribution of the type of secured creditors, it was found that non-banks and non-financial firms account for sizable percentages.

The findings in this chapter suggest several future research topics that call for more elaborate empirical analyses. On bank-firm relationships, such topics include, but are not limited to, the determinants of the duration (switching) of main bank relationships and the role of deposit relationships in corporate financing. Regarding the use of collateral, the database provides ample opportunities for deeper empirical analyses that are not possible with the datasets used in previous studies. For instance, the dataset employed here makes it possible to investigate how the role of collateral differs depending on whether it is inside or outside collateral and what determines the loan-to-value (LTV) ratio or the priority order.[23]

---

[23] Ono et al. (2014) study the cyclicality of LTV ratios and whether LTV ratios are good predictors of firms' ex-post performance.

References

Aoki, M., Patrick, H., & Sheard, P. (Eds.). (1994). *The Japanese main bank system: Its relevance for developing and transforming economies*. Oxford: Oxford University Press.

Degryse, H., Kim, M., & Ongena, S. (2009). *Microeconometrics of banking: Methods, applications, and results*. New York: Oxford University Press.

Kano, M. (2007). When does the relationship banking break up? An empirical analysis in Kansai area (relationship banking ha dono youna baai ni chuudan sarerunoka: kansai-chiiki ni okeru jisshou bunseki). In Y. Tsutsui & S. Uemura (Eds.), *Relationship banking and regional finance (relationship banking to chiiki kinyu)*. Tokyo: Nihon-Keizai Shimbun Shuppansha (in Japanese).

Kawai, M., Hashimoto, J., & Izumida, S. (1996). Japanese firms in financial distress and main banks: Analyses of interest-rate premia. *Japan and the World Economy, 8,* 175–194.

Kozuka, S., & Fujisawa, N. (2009). Old ideas die hard? An analysis of the 2004 reformation of secured transactions law in Japan and its impact on banking practices. *Thomas Jefferson Law Review, 31,* 293–315.

Mester, L. J., Nakamura, L. I., & Renault, M. (2007). Transactions accounts and loan monitoring. *Review of Finance, 20,* 529–556.

Ono, A., & Uesugi, I. (2009). Role of collateral and personal guarantees in relationship lending: Evidence from Japan's SME loan market. *Journal of Money, Credit, and Banking, 41,* 935–960.

Ono, A., Uchida, H., Kozuka, S., Hazama, M., & Uesugi I. (2010). Current status of bank-firm relationships and the use of collateral in Japan: An overview of the Teikoku Databank Data. Hitotsubashi University Research Center for Interfirm Network Working Paper Series.

Ono, A., Uesugi, I., & Yasuda, Y. (2013). Are lending relationships beneficial or harmful for public credit guarantees? Evidence from Japan's Emergency Credit Guarantee Program. *Journal of Financial Stability, 9,* 151–167.

Ono, A., Uchida H., Udell G., & Uesugi, I. (2014). Lending pro-cyclicality and macro-prudential policy: Evidence from Japanese LTV ratios. TCER Working Paper Series E-70.

Sheard, P. (1989). The main bank system and corporate monitoring and control in Japan. *Journal of Economic Behavior and Organization, 11,* 399–422.

Uchida, H., & Udell, G. (2010). Banking in Japan. In A. Berger, P. Molyneux, & J. Wilson (Eds.), *Handbook of banking, Ch. 35*. Oxford: Oxford University Press.

Uchida, H., Ono, A., Kozuka, S., Hazama, M., & Uesugi, I. (2015). *Interfirm Relationships and Trade Credit in Japan: Evidence from Micro-Data*. Springer Briefs in Economics, Tokyo: Springer

Uesugi, I., Sakai, K., & Yamashiro, G. M. (2010). The effectiveness of public credit guarantees in the Japanese loan market. *Journal of the Japanese and International Economies, 24,* 457–480.

Arito Ono is a professor at the Department of Banking and Corporate Finance, Faculty of Commerce, Chuo University, Japan. Prior to joining Chuo University in 2015, he was a senior economist at the Mizuho Research Institute, and a senior economist at the Institute for Monetary and Economic Studies, Bank of Japan (2009–2011). He has also been a member of several working groups at the Financial System Council, Financial Services Agency (2011–2015). His main fields of research are banking and corporate finance, and he has published papers in the *Journal of Money, Credit, and Banking*, the *Journal of Banking & Finance*, and the *Journal of Financial Stability*, among others. He received a B.A. in economics from the University of Tokyo and a Ph.D. in economics from Brown University.

Hirofumi Uchida is a professor of banking and finance at the Graduate School of Business Administration, Kobe University, Japan. He received his M.A. in economics in 1995 and his Ph.D. in economics in 1999, both from Osaka University. Prior to joining Kobe University in 2009, Professor Uchida was at the Kyoto Institute for Economic Research at Kyoto University and the

Faculty of Economics at Wakayama University. He was also a visiting scholar at the Kelley School of Business, Indiana University, as a 2003 Fulbright Scholar. His research has been published in the *Journal of Financial Intermediation, Economica* and the *Journal of Banking & Finance*, among others. He is also an associate editor of the *Journal of Money, Credit, and Banking*. His research interests focus on banking, financial institutions, and financial system architecture.

Souichirou Kozuka is a professor at Gakushuin University, Tokyo. He holds a Ph.D. in law from Tokyo University and had taught at Chiba University and Sophia University before beginning his teaching at Gakushuin University. While researching in his specialized field of commercial law and corporate law, he has been active in comparative law studies as well, being a correspondent of UNIDROIT, and an associate member of the International Academy of Comparative Law (IACL). He has contributed to establishing contacts with Japanese law specialists outside of Japan on many occasions and is a program convenor in eastern Japan of the Australian Network for Japanese Law (ANJeL). Dr. Kozuka serves on the editorial board of the *Journal of Japanese Law* (*Zeitschrifut für japanisches Recht*) published in Germany.

Makoto Hazama is a graduate student at Hitotsubashi University. He served as a research assistant in the project "Designing industrial and financial networks to achieve sustainable economic growth" under the program for "Promoting social science research aimed at solutions of near-future problems," conducted by the Japan Society for the Promotion of Science (2009–2013). He is working as a research assistant for the Research Institute of Economy, Trade, and Industry.

Chapter 11
What do Cash Holdings Tell us About Bank-Firm Relationships? A Case Study of Japanese Firms

Kazuo Ogawa

Abstract This chapter examines the nature of bank-firm relationships in Japan by investigating firms' cash holding behavior based on a panel dataset of Japanese firms for the 2000s provided by Teikoku Databank. This dataset has the virtue of identifying firms' main bank(s) or financial institution(s) with which they have a close relationship. This information is used to characterize the cash holding behavior of firms with varying degrees of closeness in their relationships with banks. The findings indicate that having a main bank relationship helps client firms in their cash management in two important ways. First, firms need to hold less cash for precautionary motives because main banks are ready to provide them with liquidity on a rainy day. Second, main banks can cushion shocks to client firms, so that client firms can keep the adjustment of cash holdings to such shocks to a minimum. However, client firms pay a price for maintaining long-term, stable relationships with main banks, namely, the monopoly rent imposed by main banks on their client firms in the form of a higher effective borrowing rate.

Keywords Bank-firm relationships · Main banks · Cash holdings · Precautionary saving · Monopoly rent

11.1 Introduction

Cash is held by firms for a number of reasons. In his general theory, Keynes argues that cash is held for three reasons, namely, transaction, precautionary, and speculative motives. The transaction demand for cash has been further elaborated by Baumol (1952), Tobin (1956), and Miller and Orr (1966). Since then, a number of theoretical and empirical studies have focused on the cash holding behavior of firms. Opler et al. (1999) and Bates et al. (2009) provide comprehensive surveys of the demand for cash by firms. Two further motives for a firm's cash holdings have been added to the traditional ones: tax and agency motives, as shown by Bates et al. (2009). Regarding

K. Ogawa (✉)
Institute of Social and Economic Research, Osaka University,
6-1 Mihogaoka, Ibaraki, Japan
e-mail: ogawa@iser.osaka-u.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_11

agency motives, Jensen (1986) argued that entrenched managers retain cash when a firm has poor investment opportunities. Jensen's argument is also supported by Dittmar et al. (2003), Dittmar and Mahrt-Smith (2007), and Pinkowitz et al. (2006). These authors all found that greater agency problems lead to larger cash holdings.

In Japan, agency cost problems are mitigated to a large extent by long-term, stable bank-firm relationships, known as the main bank system. A firm's main bank is frequently defined as the bank that holds the largest share of that firm's loans. However, main bank relations are not simply confined to lending relationships and cover a wide spectrum of dealings.[1] Main banks hold a large share of the loans of client firms, which gives them a strong incentive to collect information about firms' prospects and to monitor them.[2] This helps mitigate problems with information asymmetry, which can lead to adverse selection and moral hazard. Main banks also often hold both client firm debt and equity, which in terms of the agency cost approach implies that one would expect Japanese firms to hold less cash. However, Pinkowitz and Williamson (2001) found that Japanese firms in fact hold more cash than U.S. or German firms.[3] They argued that the dark side of the main bank system, namely, rent extraction, is responsible for these larger cash holdings. When a main bank exerts its monopoly power, it forces client firms to hold more cash reserves in the main bank's account. By doing so, the main bank can extract monopoly rent in the form of a higher effective borrowing rate by way of a compensating balance. Thus, two opposing forces are operating to affect cash holdings under the main bank system. This means that by examining firms' cash holding behavior, it may be possible to illuminate the nature of bank-firm relationships in Japan. This is the main purpose of this chapter.

The strategy employed to examine the nature of bank-firm relationships is to estimate a demand equation for firms' cash holdings. The estimation is based on a unique panel dataset of Japanese firms in the early 2000s provided by Teikoku Databank, Ltd. This dataset has the virtue of identifying the financial institutions with which a firm has transaction relationships in terms of loans, bills discounted, and time deposits. Using this dataset, it is possible to estimate separate demand equations for cash for different groups of firms in order to examine firms' cash holding behavior and, based on this analysis, make inferences on the nature of bank-firm relationships in Japan.

The main findings of the analysis can be summarized as follows. The results suggest that firms that have close ties with their main bank tend to hold *less* cash and the cash flow sensitivity of cash is low. This implies that main banks act as a buffer by providing liquidity to mitigate external shocks to their client firms. However, the

---

[1] Aoki et al. (1994) stressed five aspects of main bank relations: the lending relationship, client issuances of public debt, equity cross-shareholding, business settlement accounts, and the provision of information services and managerial resources.

[2] Kaplan and Minton (1994), Sheard (1994), Kang and Shivdasani (1995; 1997), Miyajima (1998), and Morck and Nakamura (1999) all provide evidence that main banks closely monitor their client firms and dispatch directors to them in the event of financial trouble.

[3] A more recent empirical study on the cash holdings of Japanese firms is Hori et al. (2010). The firms in their panel dataset are listed firms.

effective borrowing rates for these firms are significantly higher than the nominal borrowing rates and are positively related to the degree to which firms depend on their main banks (as measured by the share in total loans that the main banks account for). Thus, main banks extract monopoly rents from close client firms. In other words, having strong ties with a main bank allows firms to have lower cash holdings than would otherwise be the case but comes at the expense of a higher effective borrowing rate due to the monopoly rent extracted by main banks. In contrast, for firms with only weak ties to their main banks, cash holdings are not correlated with dependence on bank debt and the cash flow sensitivity of cash is high.

The remainder of this chapter organized as follows. Section 11.2 presents the hypotheses concerning the effects of bank-firm relationships on cash holdings within the framework of a firm's demand equation for cash. Section 11.3 then explains the dataset used for the empirical analysis and presents some descriptive statistics on cash holdings. Next, Section 11.4 presents the estimation results, which allow inferences on the nature of bank-firm relationships in Japan. Finally, Section 11.5 concludes.

11.2 Bank-Firm Relationships and Cash Holdings: Formulation of Hypotheses

Firms' relationship with their bank(s) affects firms' cash holdings in a number of ways. Specifying a firm's demand function for cash helps to clarify the channels through which such bank-firm relationships affect a firm's cash holdings. Bates et al. (2009) classified a firm's demand for cash into four motives: the transaction, precautionary, agency, and tax motive. This chapter primarily focuses on the first three motives and identifies explanatory variables corresponding to each motive.

The benchmark demand function for cash holdings is specified as follows:

$$\begin{aligned}\left(\frac{\Delta CASH}{TA}\right)_{it} &= \alpha_0 + \alpha_1 \Delta \log(SALES)_{it} + \alpha_2 log(RTA)_{i,t-1} + \alpha_3\left(\frac{\Delta NWC}{TA}\right)_{i,t} \\ &+ \alpha_4\left(\frac{CASHFLOW}{TA}\right)_{i,t} + \alpha_5 SDCFRATIO_{i,t-1} \\ &+ \alpha_6\left(\frac{DEBT}{TA}\right)_{i,t-1} + \alpha_7\left(\frac{BANKDEBT}{DEBT}\right)_{i,t-1} + \alpha_8 MAINDEP_{it} \\ &+ \alpha_9\left(\frac{CASH}{TA}\right)_{i,t-1} + \sum_{j=1}^{T}\beta_j YEARDUM_j + v_i + u_{it} \qquad (11.1)\end{aligned}$$

where v_i: firm-specific term, and
u_{it}: disturbance term.[4]

[4] Subscripts i and t represent the firm and the year, respectively.

The dependent variable is the change in cash holdings ($\Delta CASH$) divided by total assets (TA). The explanatory variables are categorized into three groups to capture each of the motives mentioned above.

11.2.1 The Transaction Motive and Bank-Firm Relationships

This section explains the variables used to represent the transaction motive of cash holdings and how the transaction motive is affected by bank-firm relationships. The growth rate of real sales ($\Delta \log (SALES)_{it}$)represents a firm's current activities as well as future investment opportunities.[5] Higher sales growth might be sustained by retaining more cash. Moreover, firms whose access to credit is constrained can use cash to make profitable investments in the future. Therefore, α_1 is expected to be positive. Evidence suggests that there are economies of scale in holding cash (e.g., Mulligan 1997). Using the logarithm of real total assets to measure firm size, it is expected that the coefficient on this (α_2) will be negative.

A change in net working capital (NWC), defined as current assets minus current liabilities minus cash, is a substitute for cash and α_3 is therefore expected to be negative. When a firm has a close relationship with its main bank, the bank will provide short-term loans when net working capital is scarce; thus, the firm does not have to keep liquidity by drawing out cash. That is, the absolute value of α_3 will be smaller for a firm with a close relationship with its main bank.

11.2.2 The Precautionary Motive and Bank-Firm Relationships

Firms will save part of their cash flow for precautionary purposes. Thus, the propensity to save α_4 will be positive. Almeida et al. (2004) demonstrated theoretically and empirically that the propensity to save is higher for financially constrained firms.[6] When the bank-firm relationship is strong, the client firm expects its main banks to provide liquidity on a rainy day, so that the firm does not need to save a lot from cash flow. Thus, the propensity to save from cash flow will be lower for a firm with close ties with its main banks.

The same argument holds for the effects of cash flow volatility on cash holdings. Han and Qiu (2007) extended the model of Almeida et al. (2004) to allow for a continuous distribution of cash flow and then theoretically showed that an increase

[5] In the literature, a widely used proxy to represent future investment opportunities is the market-to-book ratio. However, this ratio cannot be defined for unlisted firms, which make up a large part of the dataset used here.

[6] Riddick and Whited (2009) showed the opposite; they derived a negative relationship between cash flow and cash when current cash flow reveals future productivity shocks.

in the volatility of cash flow increases cash holdings for financially constrained firms.[7] This implies that there is a positive relationship between the standard deviation of cash flow to total assets (SDCFRATIO) and cash holdings ($\alpha_5 > 0$). When the relationship between a firm and its main banks is strong, the firm expects its main banks to provide liquidity in times of uncertainty; therefore, coefficient α_5 will be smaller for firms with a close relationship with their banks.

11.2.3 The Agency Motive and Bank-Firm Relationships

When a firm's debt reaches a certain level relative to equity, it faces a high risk of default and the cost of outside finance increases as a result. To avoid this situation, a debt-ridden firm will use cash to repay debt, so that the coefficient on the debt/asset ratio, α_6, is expected to be negative.[8] When the firm has a close relationship with its main bank, the main bank performs a monitoring and disciplinary role, so that the cost of financial distress will be lower and the firm does not necessarily have to pay back debt using cash. Thus, the absolute value of α_6 will be smaller for a firm with a close bank-firm relationship.

To measure a firm's dependence on banks, the ratio of debt outstanding with banks (BANKDEBT) to its total debt (DEBT) is used. When this ratio is high, the firm is likely to have a strong relationship with its bank(s). Therefore, the coefficient on the bank dependence variable, α_7, picks up the direct effect of bank-firm relationships on cash holdings. That is, a firm with strong ties to banks may hold less cash in the expectation that the banks will provide liquidity on a rainy day. In this case, α_7 will be negative. At the same time, however, when bank-firm relationships are strong, the banks may extract monopoly rents by forcing client firms to keep a large amount of cash in their accounts. In this case, α_7 will take a positive value.

Another variable employed is MAINDEP, the dependence of a firm on its main bank. The MAINDEP variable is defined as the proportion of borrowing from the main bank to total bank debt. Again, a strong main bank relationship may reduce a firm's demand for cash holdings, since the firm expects its main bank to provide liquidity on a rainy day. In this case, α_8 will be negative. On the other hand, the main bank might extract monopoly rents by forcing the client firm to keep a large amount of cash in its bank account. In this case, α_8 will positive.

The final variable used is the lagged cash/asset ratio, which measures the adjustment speed of cash holdings toward the optimal target.[9] In addition, year dummies

---

[7] Baum et al. (2008) also found that firms increase their liquidity when macroeconomic uncertainty or idiosyncratic uncertainty increase.

[8] That being said, Acharya et al. (2007) demonstrated that for constrained firms with high hedging needs the relationship between cash flow and debt as well as that between cash flow and cash holdings should be positive.

[9] Another potential determinant of the demand for cash holdings is capital expenditure. The reason that capital expenditure is not included as an explanatory variable here is that doing so would give

Table 11.1 Expected sign of cash sensitivity to its determinants under a bank-firm relationship

	Determinants of cash holdings					
	Net working capital	Cash flow	Cash flow volatility	Debt	Bank dependence	Main bank dependence
Bank-firm relationship						
Strong	−	+	+	−	−− or +	−− or +
Weak	−−	++	++	−−	−	−

A double minus or plus sign indicates that cash holdings are expected to be more sensitive to the determinant

(YEARDUM) are included in the estimation to control for macro shocks common to all firms in the sample.

Table 11.1 provides a summary of the responses of cash to its determinants and how they are affected by bank-firm relationships. Section 11.4 provides an empirical investigation of these effects.

11.2.4 Bank-Firm Relationships and Effective Borrowing Rates

When a main bank makes loans to a client firm, the firm is sometimes required to deposit part of its loans into the main bank's account. This practice raises the effective borrowing rate for the client firm. Specifically, denoting the nominal borrowing rate and the deposit rate by r_L and r_D, respectively, the effective borrowing rate (r_L^*) when a firm borrows B_L and deposits part of the borrowed money, say $D(< B_L)$, into its account with the main bank, is calculated as follows:

$$r_L^* = \frac{r_L B_L - r_D D}{B_L - D} \tag{11.2}$$

It is easy to show that $r_L^* > r_L$ as long as $r_L > r_D$. In the empirical analysis in Section 11.4, the effective borrowing rate for each firm is calculated and the correlation between these effective borrowing rates and bank-firm relationship variables is examined.

the cash holding equation more the character of an accounting identity. R&D expenditure is also frequently used as a proxy of growth opportunities. However, R&D expenditure is not available for most of the small, unlisted firms in the dataset used here.

11.3 Data Characteristics and Descriptive Statistics of Cash Holdings

This section explains the dataset used for the empirical analysis and provides descriptive statistics of cash holdings as well as major firm attributes.

11.3.1 Dataset Characteristics

The Teikoku Databank (TDB) database is a very unique and extensive dataset on Japanese firms. The dataset was constructed by a group of researchers involved in the Program for Promoting Social Science Research Aimed at Solutions of Near-Future Problems “Design of Inter-firm Network to Achieve Sustainable Economic Growth” in collaboration with Teikoku Databank, Ltd., the largest credit information provider in Japan. The dataset contains information on nearly 400,000 firms in Japan, including the financial transactions between firms and financial institutions as well as firms’ basic attributes and financial statements. A detailed explanation of the data from Teikoku Databank, Ltd. is provided in Uchida et al. (2011) and Ono et al. (2011).

The TDB dataset provides rich information on bank-firm relationships. First of all, the dataset makes it possible to identify a firm’s “main bank.” The main bank is defined as the financial institution with which the firm thinks it has the closest relationship.[10] Thus, the definition of a main bank is somewhat subjective. In addition, the amount of loans outstanding, bills discounted, and time deposits are also available for each bank-firm relationship.

The financial statements of firms are available from 2001 to 2009, although detailed information regarding bank-firm relationships is available only from 2007 to 2010.[11] Therefore, in the analysis here, firms’ cash holding behavior is examined for two periods: the whole observation period from 2001 to 2009 and the sub-period from 2007 to 2009, for which detailed information regarding bank-firm relationships is available.

For the empirical analysis, firms from non-profit-oriented industries and financial firms are excluded. Regarding the legal form of firms, the following types of firms are included in the sample: joint stock companies, closely-held limited liability companies, limited partnership companies, unlimited liability partnership companies, limited liability partnerships, medical associations, cooperative partnerships, and sole proprietorships.

---

[10] Financial institutions include both deposit-taking financial institutions and non-banks.

[11] For some firms, financial statements are available as far back as the 1990s.

Table 11.2 Descriptive statistics of firm characteristics. (Author's calculation based on Teikoku Databank data)

Year	Total assets (million yen)	Growth rate of real sales (%)	Debt-asset ratio	Bank debt/ total debt	ROA (%)
2001	227.7	1.04	0.8161	0.5687	0.64
2002	216.0	– 3.17	0.8036	0.5705	0.60
2003	206.4	– 0.26	0.7973	0.5730	0.65
2004	184.6	2.08	0.7962	0.5805	0.84
2005	187.7	3.14	0.7952	0.5782	0.83
2006	191.9	3.28	0.7920	0.5778	0.89
2007	209.5	2.99	0.7861	0.5684	0.86
2008	228.1	2.07	0.7733	0.5746	0.69
2009	397.8	– 6.27	0.7611	0.5706	0.43

11.3.2 Descriptive Statistics

Table 11.2 presents descriptive statistics of the major firm characteristics (total assets, growth rate of real sales, debt/asset ratio, ratio of short-term and long-term bank debt to total debt, and return on assets (ROA)) in terms of median values for 2001 to 2009.[12] The median value of total assets is considerably smaller than the mean. For example, the median of total assets in 2001 was 227.7 million yen, while the mean was 4,145.2 million yen. This implies that the size distribution is skewed to the right. The sales growth rate exhibits an increasing trend up to 2006 and then falls sharply in 2009. The debt/asset ratio declined gradually during the observation period from 0.8161 in 2001 to 0.7611 in 2009. Dependence on bank debt, measured by bank debt to total debt, remained rather stable during the observation period, hovering around 0.57 to 0.58. The ROA exhibits an increasing trend in the first half of the 2000s, reaching a peak in 2006 and then declining thereafter.

Table 11.3 shows the median values of the annual cash/asset ratio during the observation period. The cash/asset ratio is defined as the ratio of cash and deposits to total assets. The cash/asset ratio remained relatively stable, ranging from 0.1532 to 0.1667, during the observation period. The third and fourth columns show the cash/asset ratio by firm size. Large firms consist of firms whose total assets are larger than the sample median and small firms of firms whose total assets are smaller than the sample median. Comparing the figures in the two columns shows that the cash/asset ratio of small firms throughout the observation period is about 3–6 % points higher than that of large firms, reflecting economies of scale in cash holdings. Next, in the fifth and sixth columns, firms are divided into those whose ratio of bank debt to total debt is above and below the sample median, with the former being considered to be

[12] ROA is defined as the ratio of current net income to total assets.

Table 11.3 Descriptive statistics of cash/asset ratios classified by firm attribute. (Source: Author's calculation based on Teikoku Databank data)

Year	Whole sample	Firm size		Bank debt/total debt		Volatility of cash flow	
		Large	Small	Dependent	Not dependent	High	Low
2001	0.1550	0.1458	0.1771	0.1509	0.1630	0.1559	0.1471
2002	0.1533	0.1449	0.1722	0.1488	0.1624	0.1602	0.1390
2003	0.1532	0.1434	0.1733	0.1484	0.1627	0.1599	0.1392
2004	0.1587	0.1478	0.1781	0.1548	0.1671	0.1661	0.1433
2005	0.1627	0.1501	0.1838	0.1589	0.1700	0.1722	0.1442
2006	0.1650	0.1521	0.1880	0.1629	0.1692	0.1756	0.1438
2007	0.1667	0.1517	0.1941	0.1643	0.1713	0.1788	0.1441
2008	0.1611	0.1449	0.1915	0.1579	0.1661	0.1736	0.1388
2009	0.1622	0.1450	0.2110	0.1606	0.1643	0.1730	0.1450

bank dependent and the latter not dependent. The fifth column shows the cash/asset ratio of bank-dependent firms and the sixth column shows that of firms that are not bank dependent, and comparing the two columns shows that the cash/asset ratio is slightly smaller for bank-dependent firms. However, the difference is at most 1.2 % points, far less than the difference by firm size. Finally, the seventh and eighth columns show the cash/asset ratio dividing firms in terms of the volatility of their cash flow. Volatility of cash flow is measured in terms of the standard deviation of the ratio of cash flow to total assets over the current and past two years. Firms with a high volatility of cash flow are those whose standard deviation is larger than the sample median and firms with a low volatility are those whose standard deviation is smaller than the median. Throughout the period, the cash/asset ratio is 0.8–3.4 % points higher for firms with a higher cash flow volatility, reflecting a higher precautionary demand for cash.

Next, let us look at some descriptive statistics regarding the link between bank-firm relationships and cash/asset ratios. Figure 11.1 shows a histogram of the number of main banks in 2008. Nearly two-thirds of firms in the sample have one main bank, while 7 % of firms have two main banks. It is surprising that 27 % of the sample firms do not have any specific main bank.[13] Table 11.4 shows the cash/asset ratio for six groups of firms categorized in terms of their number of main banks and whether they are classified as bank dependent or not. The cash/asset ratio is lowest for bank-dependent firms with one main bank (0.1364), while it is highest for firms that are not bank-dependent and have no main bank (0.1770). Note that the cash/asset ratio is rather high for firms that are not bank-dependent and have more than one main bank (0.1754).

[13] The histogram for 2009 is quite similar to that for 2008, while in 2007, 74 % of the sample firms have one main bank, 11 % have two main banks and 12 % have no main bank.

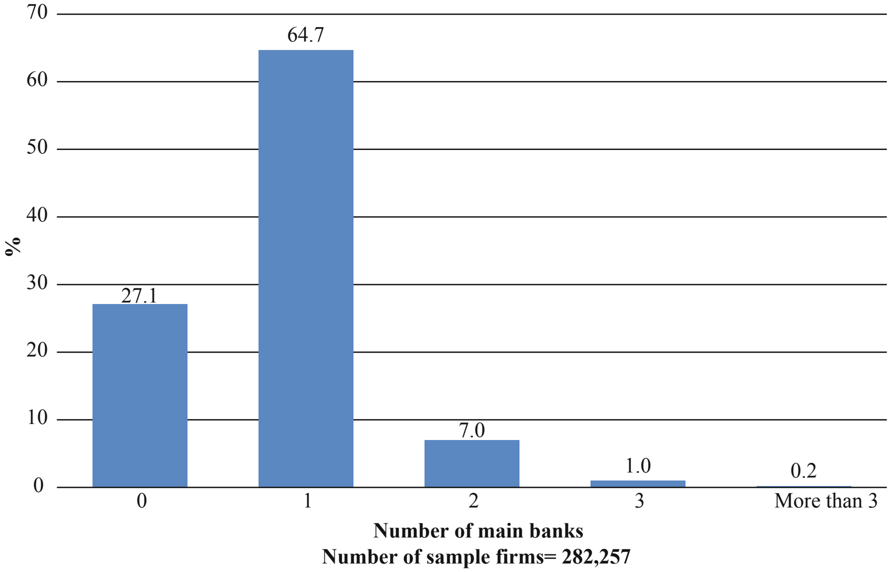


Fig. 11.1 Histogram of number of main banks: 2008

Table 11.4 Cash/asset ratios and bank-firm relationships. (Source: Author's calculation based on Teikoku Databank data)

Bank dependence	Number of main banks		
	0	1	More than one
Bank-dependent	0.1725	0.1364	0.1697
Not dependent	0.1770	0.1578	0.1754

11.4 Bank-Firm Relationships and their Effect on Firms' Cash Holdings: Empirical Evidence

Section 11.2 suggested that bank-firm relationships affect firms' cash holdings in a variety of ways. Bank-firm relationships affect not only the level of cash holdings but also the way that cash demand responds to various factors. This section provides some empirical evidence on the impact of bank-firm relationships on firms' cash holdings.

11.4.1 Estimation Results for the Whole Observation Period: 2001–2009

Table 11.5 shows the estimation results of the cash holdings equation for the whole observation period from 2001 to 2009. In the estimation, the top and bottom 1 % tails of the dependent and explanatory variables are trimmed. All explanatory variables

Table 11.5 Estimation results of cash holdings equation: baseline estimation

	Baseline estimation	Bank-firm relationship	
		Bank-dependent	Not dependent
Δ log (*SALES*)	0.0130[b] (12.08)	0.0105[b] (6.96)	0.0164[b] (10.11)
Log(RTA)	– 0.0539[b] (– 42.74)	– 0.0601[b] (– 31.93)	– 0.0428[b] (– 21.61)
NWC	– 0.2567[b] (– 100.31)	– 0.1704[b] (– 46.73)	– 0.3734[b] (– 95.73)
CASHFLOW	0.2721[b] (44.51)	0.1952[b] (21.96)	0.3602[b] (39.54)
SDCASHFLOW	0.0673[b] (5.48)	0.0394[a] (2.18)	0.1088[b] (5.79)
DEBT	– 0.1052[b] (– 23.71)	– 0.0990[b] (– 14.44)	– 0.1020[b] (– 15.39)
BANK DEBT	– 0.0358[b] (– 13.56)	– 0.0625[b] (– 10.37)	– 0.0360[b] (– 8.35)
CASH RATIO$_{-1}$	– 0.6352[b] (– 142.43)	– 0.7090[b] (– 104.18)	– 0.5566[b] (– 86.08)
CONSTANT	0.9249[b] (56.18)	1.0321[b] (41.81)	0.7603[b] (29.22)
Adjusted R-squared	0.1094	0.0791	0.1782
Number of observations	96,910	47,981	48,929
Estimation method	Fixed-effects model	Fixed-effects model	Fixed-effects model

The variables are defined as follows. Δ log (*SALES*) growth rate of real sales; *Log(RTA)* logarithm of real total assets; *NWC* change in net working capital/asset ratio; *CASHFLOW* cash flow/asset ratio;
SDCASHFLOW standard deviation of cash flow/asset ratio; *DEBT* debt/asset ratio;
BANK DEBT bank debt-total debt ratio; *CASH RATIO*$_{-1}$ lagged cash/asset ratio.
The coefficient estimates of year dummies are omitted.
Values in parentheses are t-ratios.[a] and[b] denote significance at the 5%, and 1 % level, respectively

in the baseline estimation for all firms have coefficient estimates consistent with theory and are statistically significant at the 1 % level. The level of bank debt relative to total debt (BANK DEBT) has a negative effect on cash holdings, implying that bank-dependent firms tend to hold less cash.

The next two columns of Table 11.5 report the results when estimating the cash holdings equation for bank-dependent firms and non-dependent firms separately. Again, all the coefficient estimates are significant and consistent with theory. Comparing the results in the two columns shows that cash holdings are less sensitive to net working capital, cash flow, and cash flow volatility for bank-dependent firms. This result suggests that bank-dependent firms seem to expect that financial institutions are ready to provide liquidity when needed.

To check the robustness of the results, the cash holdings equation is estimated using an instrumental variable (IV) approach. It is highly likely that the sales growth rate, net working capital, and cash flow are endogenous in the sense that common unobservable shocks simultaneously affect cash holdings and these variables. Therefore, the first-differenced cash equation is estimated using the IV method with the twice-lagged cash/asset ratio, sales growth rate, ratio of net working capital to total assets, ratio of cash flow to total assets, ratio of tangible fixed assets to total assets, ratio of intangible fixed assets to total assets, and ratio of inventory assets to total assets as instruments.[14] It turns out that some of the important variables such as the sales growth rate, the debt/asset ratio, cash flow, and the volatility of cash flow are insignificant, possibly because those variables are endogenous and the employed instruments are weak. The findings above suggested that changes in bank-dependent firms' cash holdings are less sensitive to net working capital and cash flow. Thus, this result supports the finding above that bank-dependent firms appear to expect that their bank will shield them from an external shock by providing liquidity when needed.

In the following analysis, the discussion will be based on panel estimations rather than IV estimations, which may yield estimates with large standard errors due to the weak instruments problem.

11.4.2 Estimation Results for the Sub-Period: 2007–2009

Shortening the observation period to 2007–2009 means that detailed information on bank-firm relationships such as the number of main banks and main banks' share of loans and time deposits are available, making it possible to add firms' dependence on their main bank—represented by MAINBANK and defined as the main bank's share in a firm's total loans—as a variable in the estimation. Further, it also becomes possible to split firms based on the number of main banks and their dependence on bank debt to examine how the response of cash holdings to various factors varies across firms with different degrees of closeness to their bank. Starting with the results of the baseline estimation in Table 11.6, all the coefficient estimates (with the exception of those for the sales growth rate and the volatility of the cash flow ratio) are statistically significant and consistent with theory. The results show that firms with a higher level of bank debt to total debt have lower cash holdings, which

[14] The ratio of tangible fixed assets to total assets, the ratio of intangible fixed assets to total assets, and the ratio of inventory assets to total assets are used as instruments, since they are expected to be correlated with cash flow. The estimation results are not shown to save space, but are available from the author upon request.

Table 11.6 Estimation results of cash holdings equation: 2007–2009

Explanatory variables	Baseline estimation	Bank-firm relationship	
		Bank-dependent	Not dependent
Δ log (*SALES*)	0.0028 (0.81)	– 0.0009 (– 0.20)	0.0102[a] (1.88)
Log(RTA)	– 0.1489[c] (– 21.98)	– 0.1674[c] (– 17.22)	– 0.1207[c] (– 11.32)
NWC	– 0.1932[c] (– 24.31)	– 0.0921[c] (– 8.28)	– 0.3378[c] (– 26.41)
CASHFLOW	0.2208[c] (10.27)	0.1248[c] (4.15)	0.3260[c] (9.64)
SDCASHFLOW	0.0312 (0.70)	– 0.1275[b] (– 2.14)	0.2564[c] (3.40)
DEBT	– 0.1510[c] (– 6.66)	– 0.1412[c] (– 3.96)	– 0.1450[c] (– 4.12)
BANK DEBT	– 0.0500[c] (– 4.66)	– 0.1163[c] (– 5.12)	– 0.0624[c] (– 3.65)
CASH RATIO$_{-1}$	– 1.0249[c] (– 59.91)	– 1.0772[c] (– 41.54)	– 0.9287[c] (– 37.04)
MAINDEP	– 0.0109[b] (– 2.01)	– 0.0160[a] (– 1.84)	– 0.0125[a] (– 1.75)
CONSTANT	2.3213[c] (26.84)	2.6057[c] (20.94)	1.9340[c] (14.11)
Adjusted R-squared	0.0494	0.0361	0.0866
Number of observations	22,515	11,624	10,891
Estimation method	Fixed-effects model	Fixed-effects model	Fixed-effects model

MAINDEP represents the main bank's share in a firm's total loans. For the definitions of all other variables, refer to the notes for Table 11.5. Values in parentheses are t-ratios.[a,b,] and [c] denote significance at the 10, 5, and 1% level, respectively

supports the finding obtained for the observation period as whole. In addition, firms that are more dependent on their main bank tend to have lower cash holdings.

The third and fourth columns of Table 11.6 report the results when splitting the sample into bank-dependent and non-bank-dependent firms. The results indicate that cash holdings are less sensitive to net working capital, cash flow, and cash flow volatility in the case of bank-dependent firms.[15] Further, close bank-firm relationships play an important role in allowing firms to hold less cash as they know their

[15] The coefficient estimate for cash flow volatility is significantly negative for bank-dependent firms, which is difficult to interpret.

bank will provide liquidity on a rainy day. Regarding the effect of main bank dependence on cash holdings, regardless of firms' level of bank debt to total debt, firms whose main bank accounts for a higher share of loans tend to have a lower cash/asset ratio.

To examine how the main bank relationship affects a firm's cash holdings, the sample is split into six groups of firms based on the number of main banks (none, one, and more than one) and the median of bank debt/total debt, and the cash holdings equation is then estimated separately for each group of firms. Table 11.7 reports the results, which show that the main bank relationship is closest when firms have only one main bank and the bank debt/total debt ratio is above the median. When the main bank relationship is very close, the coefficient estimates indicate that for this group of firms, cash holdings are least sensitive to the level of working capital, cash flow, and cash flow volatility. Moreover, only for this group of firms is the share of loans from the main bank associated with a significantly lower level of cash holdings. This implies that firms with close links with their main bank can keep their cash holdings at a minimum in the knowledge that if they are hit by an external shock that affects cash holdings their bank will help them out. Thus, main banks play a vital role in the liquidity management of their client firms.

Next, let us examine the cash holding behavior of firms that have the weakest ties with their main banks. These firms are the non-dependent firms with no main bank or with more than one main bank. The results indicate that the absolute value of the coefficient on net working capital is largest for non-dependent firms with no main bank, followed by non-dependent firms with more than one main bank. This implies that for these firms cash holdings and net working capital are close substitutes. The same observations hold for cash flow. The coefficient estimates for cash flow are 0.5382 and 0.4941 for non-dependent firms with no main bank and with more than one main bank, respectively. These firms save nearly half of their cash flow in the form of cash. Firms with the weakest or no main bank relationship have to rely on their own liquidity and hence their cash holdings are quite sensitive to changes in the determinants of cash demand.

Finally, it should be noted that the absolute value of the coefficient on the debt/asset ratio is largest for non-dependent firms with no main bank, reflecting that these firms have a strong incentive to use cash to redeem debt to lower the cost of external finance.

11.4.3 Monopoly Rents

As stated in Section 11.2, main banks can extract monopoly rent by requiring their client firms to deposit back part of their loans. To gauge the extent to which main banks extract monopoly rent, the effective borrowing rate is calculated in this section by taking this compensating balance practice into consideration. The monopoly rent can then be defined as the difference between the effective borrowing rate and the nominal borrowing rate. The results indicate that the monopoly rent thus calculated

Table 11.7 Estimation results of cash holdings equation with firms classified by the number of main banks and bank dependence: 2007–2009

Explanatory variables	No main bank		One main bank
	Bank-dependent	Not dependent	Bank-dependent
Δ log (*SALES*)	0.0392 (1.62)	– 0.0307 (– 1.17)	0.0037 (0.65)
Log(RTA)	– 0.2156[c] (– 5.05)	– 0.1766[c] (– 3.11)	– 0.1688[c] (– 14.74)
NWC	– 0.0931 (– 1.41)	– 0.4079[c] (– 4.15)	– 0.0848[c] (– 6.64)
CASHFLOW	0.2682 (1.43)	0.5382[c] (3.27)	0.1045[c] (3.18)
SDCASHFLOW	– 0.4211 (– 1.08)	– 0.2217 (– 0.90)	– 0.0696 (– 1.05)
DEBT	0.1154 (0.89)	– 0.3861[b] (– 2.17)	– 0.1667[c] (– 4.13)
BANK DEBT	– 0.0550 (– 0.48)	0.0715 (0.71)	– 0.0903[c] (– 3.49)
CASH RATIO-1	– 1.0266[c] (– 8.43)	– 1.1539[c] (– 8.13)	– 1.0809[c] (– 36.94)
MAINDEP	–	–	– 0.0284[b] (– 2.54)
CONSTANT	2.8213[c] (5.35)	2.7735[c] (4.62)	2.6205[c] (17.83)
Adjusted R-squared	0.0189	0.0850	0.0397
Number of observations	1208	1178	9290
Estimation method	Fixed-effects model	Fixed-effects model	Fixed-effects model
	One main bank	*More than one main bank*	
	Not dependent	*Bank-dependent*	*Not dependent*
Δ log (*SALES*)	0.0114[a] (1.88)	– 0.0166 (– 1.52)	– 0.0005 (– 0.03)
Log(RTA)	– 0.1234[c] (– 10.42)	– 0.1895[c] (– 7.57)	– 0.1378[c] (– 4.08)
NWC	– 0.3355[c] (– 23.41)	– 0.1341[c] (– 4.76)	– 0.3714[c] (– 11.55)
CASHFLOW	0.3159[c] (8.48)	0.2805[c] (3.08)	0.4941[c] (5.34)
SDCASHFLOW	0.3287[c] (3.87)	– 0.4632[b] (– 2.38)	0.4078[b] (2.05)

Table 11.7 (continued)

Explanatory variables	No main bank		One main bank
	Bank-dependent	Not dependent	Bank-dependent
DEBT	– 0.1418[c] (– 3.64)	– 0.0526 (– 0.57)	– 0.1272 (– 1.21)
BANK DEBT	– 0.0062[c] (– 3.49)	– 0.2024[c] (– 3.44)	– 0.0292 (– 0.64)
CASH $RATIO_{-1}$	– 0.9148[c] (– 31.52)	– 1.0552[c] (– 13.88)	– 0.9420[c] (– 17.24)
MAINDEP	– 0.0092 (– 1.11)	0.0041 (0.19)	– 0.0044 (– 0.24)
CONSTANT	1.9529[c] (12.87)	2.9955[c] (8.83)	2.2253[c] (5.00)
Adjusted R-squared	0.0827	0.0231	0.0941
Number of observations	8,641	1,837	1,833
Estimation method	Fixed-effects model	Fixed-effects model	Fixed-effects model

See Tables 11.5 and 11.6

is statistically significant and that, moreover, bank-dependent firms with more than one main bank relationship pay higher monopoly rent when the share of borrowing from main banks gets higher.

Specifically, the effective borrowing rate based on Eq. (11.2) is calculated using the amount of time deposits held in the main bank's accounts and the amount of loans outstanding from the main bank as collected by TDB. The nominal borrowing rate (BRATE) can be calculated from information in firms' profit-and-loss and balance sheet statements. It is assumed that deposit rates (DRATE) are the same for all sample firms but depend on the amount deposited.[16]

Figure 11.2 Presents histograms of the monopoly rent calculated for four different groups of firms classified in terms of the number of main banks (one and more than one) and their bank debt/total debt ratio.[17] The histogram of the monopoly rent of firms with a very close main bank relationship (bank-dependent firms with one man bank) resembles that of firms with a weak main bank relationship (non-dependent firms with more than one man bank). As for the descriptive statistics of the histograms, the median of monopoly rents is 27 basis points for the former group and 41 basis

[16] See the data appendix for more details on the way the nominal borrowing rate and the deposit rate are calculated.

[17] Note that we calculate the monopoly rent here by assuming that client firms are forced to hold all the time deposits. When parts of the time deposits are held at firms' own initiative, the monopoly rent would be lower.

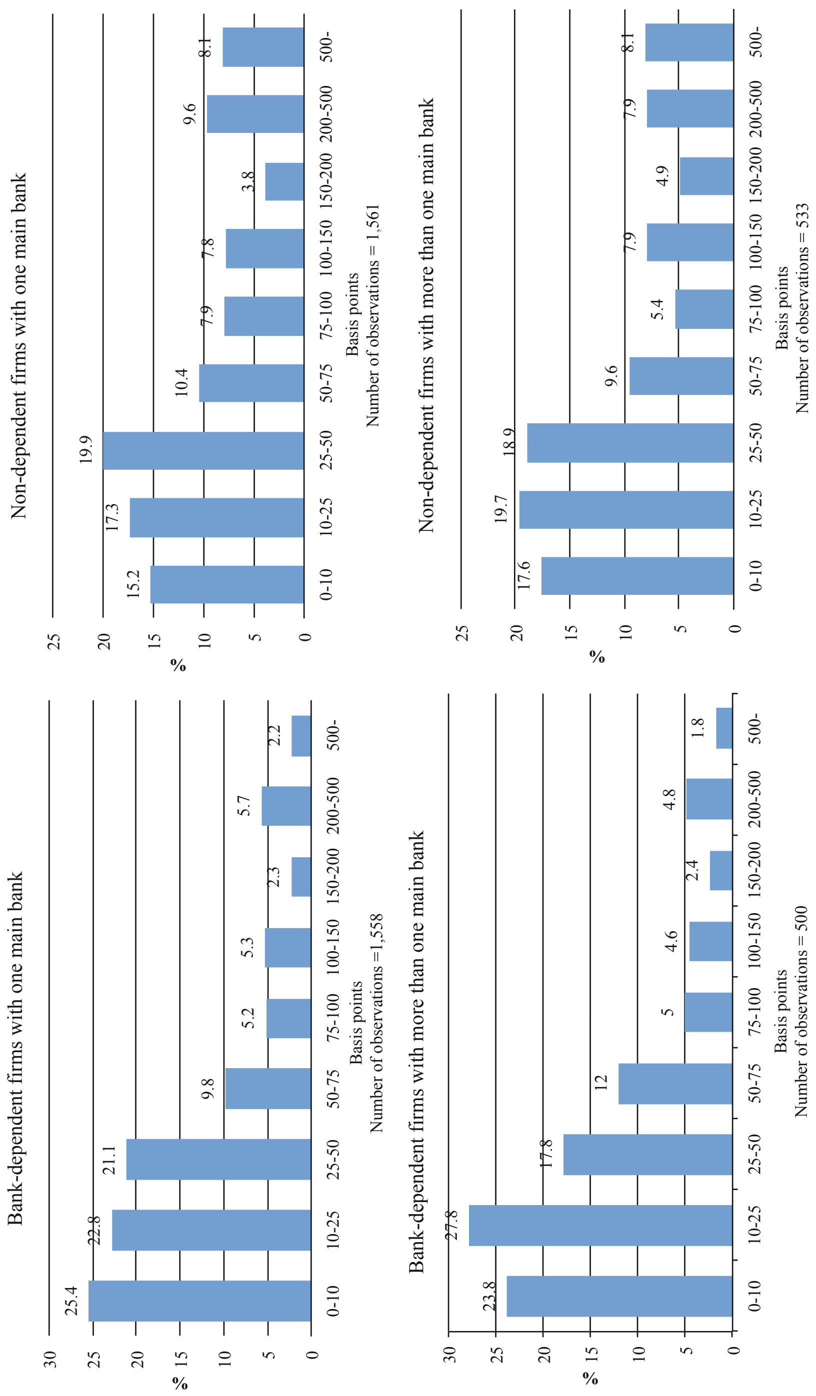


Fig. 11.2 Histograms of monopoly rent. (Source: Author's calculation based on Teikoku Databank data.)

Table 11.8 Effects of main bank dependence on monopoly rent

Firm group	Constant	Main bank dependence	Adjusted $\bar{R}^2$ Number of observations	Estimation model
Bank-dependent firms with one main bank	0.0076[c] (9.99)	– 0.0024[a] (– 1.95)	0.0032 1420	Random-effects model
Non-dependent firms with one main bank	0.0127[c] (10.73)	– 0.0021 (– 1.12)	0.0014 1401	Random-effects model
Bank-dependent firms with more than one main bank	– 0.0070 (– 1.36)	0.0193[b] (2.36)	0.0006 407	Fixed-effects model
Non-dependent firms with more than one main bank	0.0139[c] (5.34)	– 0.0050 (– 1.37)	0.0051 454	Random-effects model

Values in parentheses are t-ratios.[a,b,] and [c] denote significance at the 10, 5, and 1 % level, respectively

points for the latter. The proportion of firms for which monopoly rents are less than 50 basis points is 69.3 % for the former and 56.2 % for the latter.[18]

To further examine how monopoly rents are related to firm–bank relationships, the monopoly rent is regressed on a constant and the main bank's share of loans for each of the four groups of firms. The results are reported in Table 11.8. The constant is significant for all groups except for bank-dependent firms with more than one main bank. The results thus indicate that statistically significant monopoly rents for these three groups of firms can be detected.

On the other hand, for bank-dependent firms with more than one main bank the coefficient on the main bank's share of loans is significantly positive, which implies that the closer the main bank relationship is, the more monopoly rent the client firm has to pay.

Note that the existence of monopoly rents and lower cash holdings are not mutually exclusive. Consider, for example, the case where a firm holds a small time deposit, but the time deposit is exclusively held in its main bank's account. The main bank can then extract monopoly rent from its client firm.

11.5 Concluding Remarks

The aim of this chapter was to illuminate the nature of bank-firm relationships in Japan by looking at the cash holding behavior of firms. The findings suggest that main banks help their client firms manage liquidity in two important ways. First,

[18] The estimates of monopoly rents here are comparable to those obtained by Ono (1997). His estimates range from 20 to 80 basis points.

firms need to hold less cash for precautionary motives because main banks are ready to provide client firms with liquidity on a rainy day. Second, main banks can help to cushion unexpected shocks experienced by client firms, so that client firms can keep to a minimum the extent to which cash holdings are adjusted in response to an external shock. These are the advantages of establishing a long-term, stable relationship with a main bank.

However, client firms do pay a price for maintaining bank-firm relationships, namely the monopoly rent imposed on firms in the form of higher effective borrowing rates. Higher effective borrowing rates may lower firms' fixed investment and R&D activities, which would otherwise enhance their productivity to attain higher growth.

Thus, an important question is whether the benefits of maintaining a main bank relationship exceed the costs. This is an issue of considerable interest for future research.

Acknowledgements This research was financially supported by a Grant-in-Aid for Scientific Research (#22330068) from the Japanese Ministry of Education, Culture, Sports, Science and Technology. The author would like to thank Arito Ono, the discussant Masahiro Hori, and participants at the HIT-TDB-RIETI International Workshop on the Economics of Interfirm Networks for their helpful comments and suggestions. Makoto Hazama provided excellent assistance in constructing the dataset for the empirical analysis. Finally, the author would like to thank Teikoku Databank, Ltd. for kindly supplying the data. Any remaining errors are the sole responsibility of the author.

Appendix

This appendix explains how the variables used in the regression analysis were constructed.

1. CASH: cash and deposits.
2. TA: total assets.
3. $\Delta \log (SALES)$: growth rate of real sales. Real sales are obtained by dividing nominal sales by the GDP deflator classified by economic activity.
4. NWC: net working capital, defined as current assets minus current liabilities minus cash.
5. CASHFLOW: cash flow, measured by current net income.
6. SDCFRATIO: conditional standard deviation of the ratio of cash flow to total assets based on the current value and the value of the past 2 years.
7. DEBT: total debt.
8. BANKDEBT: short-term and long-term bank debt.
9. MAINDEP: dependence on main bank(s), measured by the ratio of borrowing from main bank (or banks) to total loans outstanding.
10. BRATE: borrowing rate defined as interest paid and discount expenses divided by the sum of short-term loans, long-term loans, bills discounted, and bonds.

11. DRATE: interest rate on time deposits. Three annual interest rates on deposits are used, namely the rate on deposits of more than 10 million yen, on deposits between 3 and 10 million yen, and on deposits of less than 3 million yen.

References

Acharya, V. A., Almeida, H., & Campello, M. (2007). Is cash negative debt? A hedging perspective on corporate financial policies. *Journal of Financial Intermediation, 16,* 515–554.

Almeida, H., Campello, M., & Weisbach, M. S. (2004). The cash flow sensitivity of cash. *Journal of Finance, 59,* 1777–1804.

Aoki, M., Patrick, H., & Sheard, P. (1994). The Japanese main bank system: An introductory overview. In M. Aoki & H. Patrick (Eds.), *The Japanese main bank system: Its relevance for developing and transforming economies* (pp. 1–50). UK: Oxford University Press.

Bates, T. W., Kahle, K. M., & Stulz, R. M. (2009). Why do U.S. firms hold so much more cash than they used to? *Journal of Finance, 64,* 1985–2012.

Baum, C. F., Caglayan, M., Stephan, A., & Talavera, O. (2008). Uncertainty determinants of corporate liquidity. *Economic Modelling, 25,* 833–849.

Baumol, W. J. (1952). The transactions demand for cash: An inventory theoretic approach. *Quarterly Journal of Economics, 66,* 545–556.

Dittmar, A., & Mahrt-Smith, J. (2007). Corporate governance and the value of cash holdings. *Journal of Financial Economics, 83,* 599–634.

Dittmar, A., Mahrt-Smith, J., & Servaes, H. (2003). International corporate governance and corporate cash holdings. *Journal of Financial and Quantitative Analysis, 38,* 111–133.

Han, S., & Qiu, J. (2007). Corporate precautionary cash holdings. *Journal of Corporate Finance, 13,* 43–57.

Hori, K., Ando, K., & Saito, M. (2010). Nippon kigyo no ryudosei sisan hoyu ni kansuru jissho kenkyu: Jojo kigyo no zaimu deta wo mochiita panel bunseki, (An empirical study on liquid asset holdings by Japanese firms: A panel analysis using financial data of listed firms). *Gendai Finance* (*Modern Finance*)*, 27,* 3–24 (in Japanese).

Jensen, M. (1986). Agency costs of free cash flow, corporate finance and takeovers. *American Economic Review, 76,* 323–329.

Kang, J. K., & Shivdasani, A. (1995). Firm performance, corporate governance and top executive turnover in Japan. *Journal of Financial Economics, 38,* 29–58.

Kang, J. K., & Shivdasani, A. (1997). Corporate restructuring during performance declines in Japan. *Journal of Financial Economics, 46,* 29–65.

Kaplan, S. N., & Minton, B. (1994). Appointments of outsiders to Japanese corporate boards: Determinants and implications for managers. *Journal of Financial Economics, 36,* 225–258.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. London: Harcourt Brace.

Miller, M. H., & Orr, D. (1966). A model of the demand for money by firms. *Quarterly Journal of Economics, 80,* 413–435.

Miyajima, H. (1998). Sengo nippon kigyo niokeru jyotai izonteki governance no shinka to henyo—Logit model niyoru keieisha kotai bunseki karano approach—(The evolution and change of contingent governance structure in the J-firm system—An approach to presidential turnover and firm performance). *Keizai Kenkyu, 49,* 97–112 (in Japanese).

Morck, R., & Nakamura, M. (1999). Banks and corporate control in Japan. *Journal of Finance, 54,* 319–339.

Mulligan, C. B. (1997). Scale economies, the value of time, and the demand for money: Longitudinal evidence from firms. *Journal of Political Economy, 105,* 1061–1079.

Ono, A. (1997). Wagakuni kinyukikan no yotai kinrizaya—tei spread no haikei to sono taiousaku-(Spread between lending rate and deposit rate in Japanese financial institutions: Background and measures against low spread). *Fuji Soken Ronsyu, IV,* 1–30 (in Japanese).

Ono, A., Uchida, H., Kozuka, S., Hazama, M., & Uesugi, I. (2011). Current status of firm-bank relationships and the use of collateral in Japan: An overview of the Teikoku databank data, Working Paper Series No. 4, Research Center for Interfirm Network, Hitotsubashi University.

Opler, T., Pinkowitz, L., Stulz, R. M., & Williamson, R. (1999). The determinants and implications of corporate cash holdings. *Journal of Financial Economics, 52,* 3–46.

Pinkowitz, L., & Williamson, R. (2001). Bank power and cash holdings: Evidence from Japan. *The Review of Financial Studies, 14,* 1059–1082.

Pinkowitz, L., Stulz, R. M., & Williamson, R. (2006). Do firms in countries with poor protection of investor rights hold more cash? *Journal of Finance, 61,* 2725–2751.

Riddick, L. A., & Whited, T. M. (2009). The corporate propensity to save. *Journal of Finance, 64,* 1729–1766.

Sheard, P. (1994). Bank executives on Japanese corporate boards. *Bank of Japan Monetary and Economic Studies, 12,* 85–121.

Tobin, J. (1956). The interest elasticity of the transactions demand for cash. *Review of Economics and Statistics, 38,* 241–247.

Uchida, H., Ono A., Kozuka, S., Hazama, M., & Uesugi, I. (2011). Current status of interfirm relationships in Japan: An overview of the Teikoku databank data, Working Paper Series No. 3, Research Center for Interfirm Network, Hitotsubashi University.

Kazuo Ogawa is professor of the Institute of Social and Economic Research of Osaka University. He also served as associate professor of the Faculty of Economics and associate professor of Graduate School of International Cooperation Studies at Kobe University (1983–1995). His research interests include macroeconomics, applied econometrics, empirical analysis of business investment, banks' behavior, and empirical analysis of households' consumption and savings behavior. Some of his major publications are *Asset Markets and Business Cycle* (awarded Nikkei Prize for Excellent Books in Economic Science), *Economics of Recession, The Truth of the Lost Decade*. He also has many articles published in leading journals on banking and the Japanese economy. He received a B.A. in economics from Kobe University in 1976 and a Ph.D. in economics at the University of Pennsylvania in 1982.

Chapter 12
Bank Lending and Firm Activities: Overcoming Identification Problems

Kaoru Hosono and Daisuke Miyakawa

Abstract This chapter presents an overview of the extant literature on the real impact of financial constraints, with a particular focus on financial constraints originating from adverse shocks to bank lending. While there has been significant progress in theoretical research on the causal link between negative fund supply shocks and various firm activities, there are relatively few empirical studies that successfully identify loan supply shocks. The first part of this chapter reviews the large body of literature on this topic and details how recent studies have attempted to overcome the important identification challenge of disentangling fund supply and demand shocks. Following the discussion of various attempts to overcome this challenge ranging from the use of natural experiments to the employment of extensive panel datasets, two studies by the authors of this chapter are discussed in detail, which employ a natural disaster in Japan as a natural experiment to examine the real impact of financial constraints on the capital investment and export behavior of firms.

Keywords Financial constraints · Bank lending · Identification · Natural experiment

---

We participated in the Study Group for Earthquake and Enterprise Dynamics (SEEDs) and the "Hitotsubashi Project on Real Estate, Financial Crisis, and Economic Dynamics" (HIT-REFINED) supported by a JSPS Grant-in-Aid Scientific Research (S), and the "Research Seminar on Corporate Finance and Firm Dynamics" held at the Research Institute of Economy, Trade and Industry (RIETI). We are thankful for financial support from Hitotsubashi University and for the data provided by Teikoku Databank, Ltd. K. Hosono gratefully acknowledges the financial support received from the Grant-in-Aid for Scientific Research (B) No. 22330098, JSPS.

K. Hosono (✉)
Department of Economics, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo, Japan
Tel.: 81-3-5992-4909
e-mail: kaoru.hosono@gakushuin.ac.jp

D. Miyakawa
Graduate School of International Corporate Strategy (ICS), Hitotsubashi University, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo, Japan
e-mail: dmiyakawa@ics.hit-u.ac.jp

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_12

12.1 Introduction

Following the seminal work of Bernanke (1983), numerous studies have investigated how shocks to bank loan supply, such as damage to bank capital or liquidity, affect the real economic activities of firms. Although empirical associations between bank loans and firm activities are frequently observed, especially during financial crises, solely from this evidence one cannot directly conclude that bank loan supply affects firm activity. As an extreme example, in frictionless complete financial markets, the financial system simply reflects changes in firms' demand for funds (see, e.g., Fama 1980; King and Plosser 1984). Put differently, any changes in bank loans merely reflect changes in loan demand and not those in loan supply. This example shows that without taking into account the fact that loan demand and supply are simultaneously determined, it is difficult to establish a causal link between fund supply shocks and firms' activities. The central challenges empirical researchers face are therefore to isolate shocks to loan supply from shocks to loan demand, and to trace the transmission of loan supply shocks to economic activities, such as the survival/exit of firms, capital investment, research and development (R&D), exports, and foreign direct investment.

The purpose of this chapter is to survey the empirical literature on bank lending and firm activities through the lens of identification strategies. While there has been significant progress in theoretical research based on asymmetric information between lenders and borrowers and incomplete contracts over the last three decades (see, among others, Bernanke and Gertler 1989; Kiyotaki and Moore 1997; Holmstrom and Tirole 1997), there are relatively few empirical studies that successfully identify loan supply shocks. Against this background, this chapter focuses on studies that are innovative in terms of identification strategies and describes them in detail. Reflecting this focus, this chapter does not necessarily provide a comprehensive survey of the entire literature.

A number of theoretical models show that if banks' capital or liquidity is impaired, they reduce loan supply, which, in turn, imposes constraints on firm activities. Bank capital or liquidity can affect economic activities through two channels. One is the bank-lending channel, through which banks with damaged capital or liquidity reduce loans more (or increase loans less) than banks with abundant capital or liquidity. The other is the firm-borrowing channel, through which borrowers find it difficult to switch without costs from banks with damaged capital or liquidity to other, undamaged lenders. This chapter primarily reviews studies that trace the impact of bank distress on both bank lending and firm activities. Note that while a number of studies examine the effects of loan supply on households (e.g., Sawada and Shimizutani 2008; Mian and Sufi 2011), the review here is restricted to research on the impact on firm activities.

The rest of the chapter is structured as follows. Section 12.2 reviews early studies on the relationship between aggregate bank lending (or bank health) and economic activities. Although a few studies have succeeded in overcoming the identification problem mentioned above, using aggregate data makes it relatively difficult to distinguish loan supply shocks from loan demand shocks. In addition, aggregate data are silent about whether shocks to banks really affect their client firms. This limitation of aggregate data naturally motivates researchers to use microdata. Section 12.3 reviews recent studies that are successful in overcoming the identification problem. Most of them use firm, bank, and bank-firm matched microdata. Next, Section 12.4 pays special attention to studies that use the 1995 Kobe earthquake as a natural experiment to identify loan supply shocks and to trace their impacts on domestic firms. Section 12.5 summarizes the review and concludes.

12.2 Early Studies on Bank Lending and Firm Activities

There are a large number of studies that empirically examine the effects of bank lending on the real economy using aggregate data. In his seminal paper, Bernanke (1983), using aggregate data, purported to show that bank failures significantly reduced aggregate production in the U.S. economy during the Great Depression. Specifically, using monthly data on industrial production, deposits at failing banks, and the liabilities of failing businesses during the interwar period, he found that the growth rate of industrial production was negatively correlated with changes in the deposits at failing banks and the liabilities of failing businesses after controlling for the unanticipated part of the rates of growth of money supply (M1) and the wholesale price index. In addition, he found that during the banking crisis the decrease in bank loans preceded the decline in output, which, according to him, suggests that changes in bank loans were driven not only by changes in loan demand due to the decline in output. Based on these findings, Bernanke (1983) concluded that bank failures and business defaults increased the cost of credit intermediation and thus affected output.

Besides bank failures, the effects of damage to bank capital on economic activity ("capital crunch") have been extensively studied. Using U.S. state-level data for the 1990–1991 recession, Bernanke and Lown (1991) found no significant relationship between bank lending and employment growth when loan growth is instrumented for by the bank capital/asset ratio, suggesting that a credit crunch was not a major cause of the 1990–1991 recession.

While many preceding studies find that damage to bank capital constrains bank loan supply, the evidence on whether firms can switch lenders without costs is mixed, possibly because the availability of alternative funding sources differs greatly across borrowers and is difficult to measure. One simple conjecture is that large and transparent firms are more likely to find alternative lenders than small and opaque firms and households. For example, Hancock and Wilcox (1998), using U.S. state-level data for 1989–1992, found that small banks reduced their loan portfolios more than large banks in response to a decline in their own capital. They also found that although reductions in the supply of credit from banks that lost capital were partly offset by other banks, they nevertheless resulted in a decline in bank lending overall. Finally, they found that economic activities (in terms of employment, payrolls, the number of firms, and gross state product) in small businesses were affected more per dollar of bank capital loss than economic activities in large businesses, although the statistical support for this hypothesis was not very strong.

12.3 Identification Strategies Used in Recent Studies

The studies discussed in the previous section have been challenged on the grounds that they do not identify loan supply shocks as distinct from shocks to loan demand. In other words, the observed association between bank failures (or financial distress) and aggregate production may simply capture the fact that both are affected by recession. To establish a causal linkage between loan supply shocks and firm activities, it is necessary to identify cases of bank failure or financial distress that have little to do with the level of local economic activity. To confirm and quantify the effect of bank loan supply on firm activities, recent studies have developed a number of identification strategies to isolate shocks to the supply of bank loans from shocks to the demand for bank loans, and have traced the transmission of these shocks to bank lending and further to firm activities. Before looking at individual studies in detail, it is useful to briefly explain and classify these identification strategies.

One approach is to conduct an event study that examines the stock market performance of firms that have a relationship with a failed bank. Using high-frequency (daily) data, one has a good chance to identify shocks to bank loans associated with bank failures and estimate their impact on client firms' share prices. Slovin et al. (1993), Yamori (1999), Yamori and Murakami (1999), Bae et al. (2002), and Brewer et al. (2003a, b) follow this strategy. As will be discussed below, however, this approach has a number of serious limitations. To overcome the limitations of the use of stock market performance, a number of studies use information on financial statements of firms to examine the impact of a bank failure on client firms, although these studies suffer from serious endogeneity problems.

Another approach is to make use of geographical or economic borders. Suppose that banks in economy A suffer damage to their capital or liquidity due to a shock inside A, such as the collapse of a bubble in domestic asset markets, a natural disaster, or new regulations. Under these circumstances, the financial difficulties of banks in A are unlikely to be related to changes in loan demand by firms in economy B. Therefore, changes in the supply of loans by banks in A can be considered to be completely exogenous to firms in B. Suppose further that while banks in A make loans to firms in B, firms in A do not transact with firms in B (i.e., there are no real linkages through trade or customer-supplier relationships). In this case, it is possible to clearly trace the transmission of a shock in bank loan supply in A to firm activities in B. The first study to employ this line of reasoning and examine the international transmission of a financial crisis in one country on economic activities in another using aggregate data is Peek and Rosengren (2000). Other studies following a similar strategy using aggregate data are van Rijckeghem and Weder (2001) and Chor and Manova (2012).

The identification strategies employed in Peek and Rosengren (2000) and subsequent studies did address the identification problem. Nonetheless, the use of aggregate data still makes it difficult to determine whether only firms borrowing from the damaged banks or firms in general were affected. To examine this issue more rigorously, one needs to utilize information on the relationship between firms and banks. Studies doing just that include those by Chava and Purnanandam (2011), Paravisini et al. (2011), and Schnabl (2012), which investigated the international transmission of financial crisis using matched bank-firm data. In contrast with these studies on the international transmission of loan supply shocks, there are relatively few studies on the domestic transmission of loan supply shocks, with exceptions being Hosono et al. (2012) and Miyakawa et al. (2014).

Yet another identification strategy employed in recent studies is to employ information on loans from individual banks to firms that borrow from multiple banks. Suppose that a firm's total amount of borrowing increases but that from one particular bank decreases. In this case, one can reasonably regard the decrease in borrowing from that one bank as reflecting a decrease in loan supply from that bank rather than a decrease in loan demand. This line of reasoning points to an empirical strategy that attempts to identify (bank-specific) loan supply shocks based on changes in loans from a particular bank to a particular firm. In most studies employing this strategy, changes in individual loans are regressed on a measure of loan supply shocks controlling for firm-level fixed effects, which account for changes in firm-specific loan demand. The first to employ this approach were Khwaja and Mian (2008) using loan-level data to identify loan supply shocks. Jiménez et al. (2012) and Amiti and Weinstein (2013) have further elaborated this approach.

Given that it is generally difficult to obtain detailed firm-bank matched data on borrowing, a number of recent studies have utilized survey data containing information on whether a firm has applied for a loan and whether this application has been accepted by the bank. Although such data are useful for measuring firms' loan demand, it is not straightforward to use these data to analyze loan supply shocks. The reason is that a rejection of a loan application can reflect both shocks to loan supply and demand. In other words, it is still necessary to identify supply shocks using a clear identification strategy. Suppose that the proportion of rejections among applications increases in the wake of an external shock to the banking sector such as a financial crisis. In this case, the increase in rejection rates can be regarded as a negative shock to loan supply only if it is possible to ascertain that the shock is truly exogenous to the banking sector. Studies using survey data to examine loan supply shock include those by Popov and Udell (2010), Puri et al. (2011), and Minetti and Zhu (2011). How they utilize the survey data using various empirical methodologies will be discussed in detail later in the chapter.

There are also a number of other identification strategies. Driscoll (2004), for example, regarded the inflow and outflow of deposits due to local demand for liquidity in one state as a shock to loan supply in another state. On the other hand, Ashcraft (2005) focused on the effect of the failure of healthy banks that are subsidiaries of a failed multi-bank holding company, while Amiti and Weinstein (2011) used variations in bank health variables that are uncorrelated with changes in a firm's share price as an exogenous shock to bank loans. Each of these identification strategies and the results of the studies employing them are described in more detail below.

12.3.1 Event Studies and Bank Failure

One strand of studies follows the event study approach to see how the stock prices of firms change when their lending banks fall into financial trouble. Slovin et al. (1993) were the first to analyze the stock prices of firms that had lending relationships with Continental Illinois Bank during the period of its de facto failure. They found that after adjusting for overall stock price changes, the stock prices of Continental Illinois client firms on average fell by 4.2 % during the bank's impending insolvency but then gained 2.0 % on average in response to the announcement that the government would rescue the bank. They concluded that bank financial distress harms client firms as a result of their loss of relationship-based cost advantages intrinsic to bank lending. The study by Slovin et al. (1993) was followed by similar ones by Yamori (1999), Yamori and Murakami (1999), Bae et al. (2002), and Brewer et al. (2003a, b); all found that bank failures had a significant effect on the market value of firms that borrowed from those banks.[1]

The advantage of these event studies is that they are able to clearly identify bank failure shocks using high-frequency (daily) data. Given that loan demand is unlikely to change within a few days subsequent to a bank failure, high-frequency data should contain information on the effects of loan supply shocks that originate from bank financial distress on borrowers' performance. Nonetheless, such event studies have limitations as well. First, event studies rely on the assumptions of market efficiency and rational investor behavior, which have been challenged in various studies. Second, event studies cannot be applied to non-listed firms, for which bank lending is likely to matter more than for listed firms. The studies discussed in detail below, including two recent ones conducted by the authors of this chapter, focus on the real activities of firms and hence do not require any assumptions of market efficiency or rationality. In addition, some of these studies analyze unlisted firms, most of which are small and medium-sized and are therefore likely to be severely affected by shocks from lending banks.

Given the limitations of the event study approach, several studies employ information on firms' financial characteristics and on firm-bank relationships to investigate the effects of bank failures or weak bank health on client firm performance in terms of profitability, investment, exports, or survival. As discussed above, this approach is valid only if the bank failure is exogenous to firm performance. Unless this is the case, the analysis faces exactly the same endogeneity problem already mentioned. Nonetheless, it is worthwhile reviewing these studies as they are closely related to the event studies mentioned above.

For instance, Hori (2005) examined the profitability of firms that borrowed from a large failed Japanese bank and found adverse effects on firms with low credit ratings. Similarly, Minamihashi (2011) analyzed the failures of two long-term credit banks

[1] Note, however, that Brewer et al. (2003a) also found that the magnitude of these negative effects on the value of borrower firms was not significantly different from that on all other firms in their sample.

in Japan and found that the failures significantly decreased the investment of their client firms (see also Fukuda and Koibuchi 2007). In addition to bank failures, the effects of damage to bank capital on borrowing firms' investment behavior have also been explored in several studies. For example, Gibson (1995, 1997) found that client firms that borrowed from Japanese banks with low credit ratings significantly reduced their investment during 1994–1995.[2] As already mentioned, it should be noted that these studies could still suffer from the endogeneity problem, because poor firm performance may have caused bank failures or financial distress.

12.3.2 Use of Geographical Borders

12.3.2.1 Evidence of International Transmission Using Aggregate Data

To overcome the identification problem, Peek and Rosengren (2000) used the Japanese banking crisis as a natural experiment and investigated whether loan supply shocks can affect real economic activities. Notably, Japanese banks, which were severely hit by the decline in equity and commercial real estate prices in Japan in the early 1990s, suffered a deterioration in their capital base and reduced their lending in the United States. Because the shock was external to U.S. credit markets but affected them through the substantial penetration of the U.S. commercial real estate loan market by Japanese banks, the authors were able to use this event to identify an exogenous loan supply shock and to link that shock to construction activity in major U.S. commercial real estate markets. More precisely, as a first step, they regressed the change in bank-level commercial real estate loans in certain U.S. states on variables representing the link of Japanese bank subsidiaries in the United States with their parent in Japan and the health of the parent bank while controlling for other covariates potentially affecting bank lending.[3] The regression results clearly show the international transmission of the poor financial condition of Japanese parent banks to local subsidiaries. As a second step, they further examined whether aggregate state-level construction activity was negatively affected by the international transmission of the Japanese banking crisis and found that this was the case. Their findings imply that the retrenchment of Japanese lending had a substantial adverse impact on U.S. construction activity, indicating that at least some borrowers were not able to easily obtain alternative financing.

A number of empirical studies following Peek and Rosengren (2000) used the same identification strategy. First, Van Rijckeghem and Weder (2001) examined the

[2] Using matched bank-firm data for Japan from 1987 to 1994, Kline et al. (2002) found that financial difficulties of banks significantly reduced the number of foreign direct investment projects by Japanese firms in the United States. Using data of listed Japanese firms for 1993–1999, Peek and Rosengren (2005) found that banks expanded loans to unprofitable firms during this period. See also Caballero et al. (2008) for such "zombie" lending practices by Japanese banks in the 1990s.

[3] For this part of their analysis, they used bank-level (i.e., micro) data.

potential contagion of the Mexican, Thai, and Russian crises to other countries. They constructed several measures accounting for the similarity between the "ground-zero" countries (i.e., Mexico, Thailand, and Russia) and other countries in terms of the patterns of international trade and finance. Their conjecture was that countries are more likely to face an adverse impact through the financial channel if they borrow from lender countries that happen to be lenders to the ground-zero countries. Using country-level international trade and finance data, they confirmed the existence of such a common lender effect (i.e., the financial channel) while controlling for real linkages (i.e., the trade channel). Second, Chor and Manova (2012) studied the impact of the global financial crisis of 2008–2009 on international trade flows. Their empirical methodology involved regressing U.S. imports from each country-sector pair in each year on various independent variables including (i) country-year specific interbank rates, (ii) the sector specific dependence on external finance, (iii) a dummy variable taking a value of 1 for years since the start of the financial crisis, and (iv) a triple interaction term among these three variables. Using monthly international trade data, they found that countries with tighter financial conditions (proxied for by higher interbank rates) exported less to the United States during the 2008–2009 global financial crisis. They also found that this pattern became more apparent in sectors showing a higher dependence on external finance, illustrating that financial constraints have an impact on real economic activity.

12.3.2.2 Evidence of International Transmission Using Matched Bank-Firm Data

As stated previously, although the identification strategies employed in Peek and Rosengren (2000) and other studies were successful in overcoming the identification problem, the use of aggregate data makes it difficult to determine whether only firms borrowing from the damaged banks or firms in general were affected. Moreover, two further concerns have emerged in this strand of literature. First, aggregate data do not take into account individual-level heterogeneity. For instance, the adverse effects of a loan supply shock may be more severe for firms with a weak balance sheet than for other firms. In this regard, the natural response is to use individual-level (e.g., firm-level and matched bank-firm level) data to directly address this concern. Second, the loan supply shock itself could be measured at an individual bank level. While Peek and Rosengren (2000) measured the loan supply shock at the state level, it is possible to measure the shock at the bank level by using, for example, data on each bank's exposure to mortgage-backed securities that led to huge losses during the recent financial crisis. These considerations mean that it is desirable to use more detailed (e.g., bank-level) information on shocks to the supply of funds.

Chava and Purnanandam (2011), Paravisini et al. (2011), and Schnabl (2012) addressed these issues using matched bank-firm data. First, Chava and Purnanandam (2011) measured firm heterogeneity in terms of firms' bank dependence and examined whether bank-dependent firms were more adversely affected by the shock to the U.S. banking system originating from the Russian crisis in 1998. They found that bank-dependent firms faced a larger decline in their firm value than firms that issued bonds in the past. To further take the heterogeneity of loan supply shocks into account, Chava and Purnanandam (2011) constructed a bank-firm match-level dataset using the DealScan database to identify each firm's main bank. To measure the extent to which these main banks were exposed to the Russian crisis, they used information in the quarterly call reports filed by commercial banks. Using the bank-firm match-level data augmented by each main bank's heterogeneous exposure to the crisis, the authors found that, in the period following the crisis, crisis-affected banks reduced the quantity of lending more and increased interest rates charged to their clients more than unaffected banks.

Second, another example of a study that succeeds both in taking firm-level heterogeneity into account and employing individually measured financial shocks is that by Paravisini et al. (2011), who used Peruvian bank-firm match-level data with customs information about the exported products of each firm and examined whether the reversal of capital flows during the 2008 global financial crisis had a greater adverse impact on both the extensive margin of exports (i.e., the number of firms continuing exports) and the intensive margin of exports (i.e., the volume of exports) for firms that had borrowed more from banks that were more exposed to the financial crisis. To take into account the possible endogeneity of fund supply, they instrumented for the supply of credit to a firm using an exogenous shock (i.e., the reversal of capital flows) affecting the lender bank's balance sheet. Their results suggest that a negative credit supply shock significantly reduces both the extensive and the intensive margins of exports.

Third, also using firm, bank, and bank-firm match-level data for Peru, Schnabl (2012) examined the liquidity shock originating from the 1998 Russian sovereign default. One of the key features of his study is that he investigates the impact of the Russian default on (i) the lending from international banks to Peruvian banks, (ii) the lending from Peruvian banks to Peruvian firms, and (iii) real measures of the performance of Peruvian firms (e.g., loan defaults and firm survival). Like Khwaja and Mian (2008), whose study will be discussed in detail later, Schnabl (2012) controlled for bank- or firm-level individual effects in changes in fund demand when conducting his estimations to examine issues (i) and (ii) above. Including these fixed effects, he was able to identify that causality ran from the Russian default to (i) and (ii). However, he did not control for firm fixed effects when estimating (iii), since the unit of observation in that estimation is the firm rather than the loan relationship. His results showed a significant effect for (i) and (ii). As for (iii), interestingly, he did not find a monotonic relationship between the liquidity shock to banks and firms' performance. While firms borrowing from banks with low or intermediate exposure to the impact of the Russian default showed better performance than firms that borrowed from banks with high exposure, firms borrowing from banks with intermediate exposure showed better performance than firms that borrowed from banks with low exposure. He conjectured that firms borrowing from banks with low exposure to the Russian default may have had lower credit demand than firms borrowing from banks with intermediate exposure. For firms with lower credit demand, a liquidity shock to banks matters less for their performance than for firms with higher credit demand. He showed that banks with low exposure to the Russian default actually lent to smaller firms which have lower credit demand and found that the impact of the liquidity shock on firm performance is misestimated due to the lack of firm fixed effects in the regression for (iii).

Although matched bank-firm data are useful for identifying loan supply shocks, another issue may arise, namely, endogeneity in the matching of firms and banks. Suppose that banks specialize in a certain class of firms and that such class of firms happens to be hit by a decline in demand for a certain product or in a country that the firms export to. Then such banks are affected by the product- or country-level demand shock, and this may shrink the supply of loans. At the same time, firms' demand for loans may decrease due to the same shock. To overcome the bias originating from this potential endogeneity in the matching, it is useful to control for product- and country-level fixed effects. Paravisini et al. (2011) controlled for such fixed effects at the firm-product-destination level to exclude this type of matching mechanism when they estimated the effects of credit supply on firm exports.[4]

12.3.2.3 Evidence of Domestic Transmission Using Matched Bank-Firm Data

While there are a considerable number of studies focusing on the transmission of shocks in an international context to overcome the identification problem, there are only a few studies that succeeded in examining the domestic transmission of shocks. The reason is that in a domestic setting it is difficult to find appropriate proxies for fund supply shocks that are independent from fund demand shocks. Against this background, a number of studies have used natural disasters (Hosono et al. 2012; Miyakawa et al. 2014) or man-made events (Khawaja and Mian 2008) as natural experiments. For example, the identification strategy adopted by Hosono et al. (2012) and Miyakawa et al. (2014) is based on the reasoning that a large earthquake may cause severe damage to enterprises and banks in the area without affecting enterprises and banks located far away from the earthquake area. This kind of setting provides a natural experiment that makes it possible to compare firms not affected by the earthquake transacting with unaffected banks and unaffected firms transacting with affected banks. In other words, the setting resembles those examined in the literature on the international transmission of financial shocks. Section 12.4 below discusses Hosono et al. (2012) and Miyakawa et al. (2014) in detail.

[4] The study by Amiti and Weinstein (2011), which is described in Section 12.3.5 below, deals with this problem of endogenous matching by adding a full set of bank fixed effects. Specifically, they add a dummy variable indicating whether or not the firm is the bank's client in that year and its interaction term with the change in banks' market-to-book value.

12.3.3 *Use of Loan-Level Information for Firms with Multiple Bank Relationships*

Data on firms' borrowing from multiple banks can be used to identify changes in loan supply. A prime example is the study by Khawaja and Mian (2008). Identifying a (bank-specific) loan supply shock as a change in loans after controlling for firm-level fixed effects, which reflect firm-specific loan demand shocks (as well as aggregate loan supply shocks), they used the large-scale withdrawal of foreign currency (dollar) deposits in Pakistan immediately after the suspension of exchange rate liquidity support by the International Monetary Fund (IMF) in response to the testing of a nuclear device in India and Pakistan as a natural experiment. Owing to the partial freeze of dollar deposits declared by the prime minister of Pakistan in the wake of the IMF step, Pakistan's banking sector encountered a serious negative shock. Using matched bank-firm loan data and information on the characteristics of each bank, including changes in deposits, Khawaja and Mian (2008) tested whether there was a relationship between the size of withdrawals that a firm's lender banks experienced and the decline in lending to the banks' client firms. The key result of their analysis is that banks did pass on the negative shock due to the reduction in their deposits to their client firms. In addition, by regressing changes in the rate of firm defaults on the average change in lender banks' deposits, they found that the transmission of the shock had a direct impact on firm default rates. Interestingly, they further found that for larger firms, no adverse impact on the default rate was found, presumably because such firms were able to more easily switch to banks facing smaller withdrawals of deposits than smaller firms.

Using a similar approach, Jiménez et al. (2012) used data on loan applications in Spain to examine how changes in aggregate variables such as interest rates and GDP as well as the interaction between these variables and lender bank characteristics affect the likelihood of loan applications being granted. Extending the empirical strategy used in Khwaja and Mian (2008), they controlled for the time-variant quality of potential borrowers by considering either firm-month or loan-level fixed effects.[5] They found that higher short-term interest rates and lower GDP growth reduced the probability that a loan application was granted and that this tendency was stronger for banks with low capital (in periods of higher short-term interest rates and lower GDP growth) or low liquidity (in periods of higher short-term interest rates). They also found that firms whose initial loan application was rejected were unable to reverse the resultant reduction in credit availability by applying to other banks, especially in periods of tighter monetary and economic conditions.[6]

[5] They utilized information on a firm's successive loan applications to different banks when they controlled for loan-level fixed effects.

[6] Hosono and Miyakawa (2014) applied a similar identification strategy to data on outstanding loans for bank-firm pairs in Japan and found that banks with more liquidity or capital tended to lend more to their client firms. Moreover, the size of these effects tended to be larger when economic growth was lower. The study further found that the transmission of the shock had a direct impact on firms' capital investment.

12.3.4 Use of Information on Loan Applications and Acceptance/Rejection

This subsection reviews studies that use survey data regarding firms' loan applications and their success or failure. As already mentioned, using such survey data does not necessarily help to overcome the identification problem. The goal in this section is to detail what empirical techniques and survey data have been used to identify loan supply shocks.

12.3.4.1 Survey Evidence on International Transmission

To examine the international transmission of financial crises, Popov and Udell (2010) and Puri et al. (2011) used survey data containing information on whether a firm applied for a bank loan and whether the application was successful.

Specifically, starting with the former study, Popov and Udell (2010), using data on local small and medium-sized enterprise financing in 14 central and eastern European countries, examined the effect of the financial distress of the foreign parent banks during the early stage of the 2007–2008 financial crisis on small firms through the local branches or subsidiaries of the foreign banks. In this sense, they employ the same identification strategy as in the literature pertaining to geographical borders. Combining unique survey data on (i) whether firms were seeking to take out a loan, (ii) firms had actually applied for a loan, and (iii) the results of firms' loan applications with information on city-level foreign bank penetration measures, which is necessary to link the bank-level information and firm-level information at the city level, they showed that firms located in cities with a higher penetration of foreign banks affected by the 2007–2008 financial crisis faced a severer credit crunch. The finding implies that foreign banks transmitted the shock originating from the financial crisis to local borrowers.[7]

Another example of the use of survey data on the cross-border transmission of financial shocks is the study by Puri et al. (2011), who employed retail banking data for Germany, including information on the acceptance and rejection of loan applications, to test whether local saving banks exposed to the 2007 U.S. financial crisis transmitted the shock to their customers. In addition to examining whether and how the impact of the crisis spread from banks directly affected by the crisis, they were also interested in whether and how lending relationships between customers and banks before the crisis mitigated the adverse impact of the crisis. Their findings show that the negative shock to fund supply significantly reduced the probability that customers' loan application was accepted; however, this effect was mitigated where banks had fewer liquidity constraints and pre-crisis customer-bank relationships existed.

[7] Popov and Udell (2010) did not use observations on actual bank-firm matches, but instead matched firms and banks based on their locality.

12.3.4.2 Survey Evidence on Domestic Transmission

Minetti and Zhu (2011) also used information on loan applications from firm surveys, but they examined the domestic transmission of loan supply shocks. Specifically, using an Italian firm-level survey on firms' subjective assessment of the extent to which they are credit constrained, they identified financially constrained firms and the impact of these constraints on firms' exports. They also addressed two potential endogeneity issues. The first potential endogeneity issue relates to the fact that whether a firm is rationed may be correlated with firm attributes such as firms' productivity that are observable only to lenders but not to the researcher. The second potential source of endogeneity is that firms with severe agency problems may be more exposed to rationing and, at the same time, more likely to export (for the purpose of, say, "empire building"). Regressing firms' export activities on their survey response on credit rationing and control variables may still lead to the misestimation of the effect of credit rationing. Thus, to overcome these endogeneity issues, Minetti and Zhu (2011) employed a strategy that employs an exogenous restriction on the local supply of banking services. Specifically, they used the banking regulations implemented in each Italian region in the past as instruments based on the assumption that such past regulations affect firms' current export behavior only through the financial constraints originating from lender banks' current loan provisions. The past regulations affected bank location choice and competitiveness, which eventually affected the degree of financial frictions faced by firms located in each region. As the implementation of past bank regulations is completely exogenous to firms, the approach should be able to identify the impact of financial frictions on firm activities. Their results show that both the extensive margin of exports (i.e., the probability of exporting) and the intensive margin of exports (i.e., the amount of foreign sales) are significantly lower for rationed firms.

12.3.5 Other Identification Strategies

A number of alternative identification strategies can be found in the literature. First, Driscoll (2004) exploited the fact that U.S. states can be viewed as a group of small open economies. This means that state-specific shocks to money demand are automatically accommodated, leading to changes in lending if banks rely on deposits as a source of lending. For example, an increase in demand for deposits in one state results in a flow of funds from other states and an increase in deposits in that state. Thus, the supply of loans increases in that state. Based on this idea, Driscoll used state-specific money demand shocks as instruments for loan supply when regressing output on loans. Using a panel of annual data on U.S. states, he found that although shocks to money demand have a significant effect on loans, upon instrumentation, shocks to the supply of loans do not have a significant effect on state personal income.

Second, Ashcraft (2005) examined the effects of bank failures on local economic activity by identifying bank failures that occurred for reasons that have little to do with local economic activity. Specifically, using a county-level dataset from the United States, he studied two similar incidents that occurred when the healthy subsidiaries of two multi-bank holding companies collapsed after the failure of their unhealthy lead banks. In both events, the failure of the unhealthy lead bank led to the failure of the healthy subsidiaries, since the Federal Deposit Insurance Corporation (FDIC) applied cross-guarantees. Ashcraft (2005) found that the failure of the healthy subsidiaries due to the failure of the unhealthy lead banks had significant and persistent effects on real economic activity in both cases.

Finally, using matched bank-firm data on Japanese firms and banks providing them with trade finance from 1990 to 2010, Amiti and Weinstein (2011) regressed firms' export growth rates on variables representing banks' health. To check the robustness of their results, they addressed the endogeneity problem that may arise from ordinary least squares (OLS) estimation by using the residuals from a regression of changes in banks' health variable (i.e., the market-to-book value) on changes in firms' share price as an instrument. They concluded that a deterioration in banks' health results in a significant decrease in client firms' exports relative to output, suggesting that trade finance does play a role as a determinant of exports.

12.4 Two Empirical Studies Using an Earthquake as a Natural Experiment

This section provides details of two studies conducted by the authors of this chapter (Hosono et al. 2012; Miyakawa et al. 2014) using a major natural disaster as a natural experiment to examine the transmission mechanism of a financial shock in a domestic setting. Specifically, both studies use the Great Hanshin-Awaji (Kobe) Earthquake (which hit the areas around Kobe City and Awaji Island in western Japan in January 1995) as an exogenous shock to Japanese banks. The studies examine whether damage to banks had an adverse impact on the capital investment and exports of client firms that did not themselves suffer any damage. The studies employ a unique firm-level dataset compiled from various sources especially for this purpose. The dataset contains data on firms' investment and export activities as well as information on whether banks and firms were located inside or outside the earthquake-affected area. By comparing the investment and export behavior of undamaged firms borrowing from damaged banks with that of undamaged firms borrowing from undamaged banks, it was possible to identify the effect of damage to banks on these two important firm activities.

12.4.1 Capital Investment

In order to examine the impact of the loan supply shock resulting from the natural disaster on firms' capital investment, Hosono et al. (2012) estimate the following investment equation, which is based on Tobin's q and augmented by a dummy variable indicating whether a firm is located in the earthquake-affected area as a proxy for firm damage, a proxy for bank damage which we describe below, proxies for a firm's financial constraints, and proxies for the lending capacity of the firm's main lending bank:

$$\begin{aligned}\frac{I_{it}}{K_{it-1}} = \beta_0 &+ \beta_1 F\_SALESGROWTH_{it-1} + \beta_2 F\_DAMAGED_i \\ &+ \beta_3 B\_DAMAGED_i \\ &+ \beta_4 F\_DAMAGED_i * B\_DAMAGED_i \\ &+ \beta_5 F\_CONSTRAINTS_{i,t-1} \\ &+ \beta_6 B\_CAPACITY_{it-1} + \beta_7 Industry_i + \varepsilon_{it}, \\ &\mathit{for}\ t = 1995,\ 1996,\ 1997. \end{aligned} \tag{12.1}$$

The dependent variable is the capital investment ratio, which is defined as the ratio of investment during period t to the capital stock at the end of period $t-1$. Taking into account the possibility that the effects of the earthquake on investment change over time, separate cross-sectional regressions for each fiscal year are run.[8]

The explanatory variables include firms' sales growth as a proxy for Tobin's q as well as a number of further variables that may affect investment. Specifically, we include *F_DAMAGED*, which takes a value of 1 if the firm is located in the earthquake-affected area. The main interest of the analysis lies in the effects of bank damage on borrowing firms' investment, which in the regression is picked up by the variable *B_DAMAGED*, a variable representing damage to a firm's main bank. In fact, because no precise information on whether and to what extent banks suffered damage from the earthquake is available, two alternative variables are used for *B_DAMAGED*. The first is *B_HQDAMAGED*, a dummy variable that takes a value of 1 if the headquarters of the bank was located in the earthquake-affected area. This variable captures whether or not the managerial capacity to process loans was impaired; this managerial capacity includes back-office operations such as the ability to process applications for large loans or to manage the total risk of the bank's loan portfolio.

The second variable used is *B_BRDAMAGED*, which is the share of the main bank's branches located in the earthquake-affected area as a fraction of the total number of branches. Thus, this variable measures the extent of damage to the main bank's branch network. It represents the impairment of the main bank's ability to process applications for relatively small loans under the authority of branch managers. It also captures the extent of the main banks' exposure to damaged and possibly non-performing borrowers, which was likely to negatively affect their risk-taking

[8] Fiscal year t begins in April of year t and ends in March of year $t+1$.

capacity. The studies hypothesize that these measures of bank damage capture borrowing constraints on client firms and are therefore expected to take a negative coefficient in the regressions.

Note that firms' main banks are defined as of the time that the earthquake occurred (i.e., in FY1994) so as to properly identify an exogenous shock to the firm, i.e., whether or not the firm's main bank at the time of the earthquake sustained damage. If firms can easily switch their main bank, this would allow them to escape any adverse effects resulting from damage to their main bank, in which case the size of the coefficient on *B_HQDAMAGED* and *B_BRDAMAGED* would likely be small.

In addition to *F_DAMAGED* and *B_DAMAGED*, an interaction term of these two variables is added in order to differentiate the impact of bank damage on damaged firms from that on undamaged firms. As mentioned earlier, what the studies are most interested in is the effect of bank damage on undamaged borrowers, which is captured by the coefficient on *B_DAMAGED*.

The regressions further include a vector of variables representing a firm's financial constraints, *F_CONSTRAINTS*, consisting of the firm's size (the natural logarithm of total assets), its leverage (the ratio of total liabilities to total assets), profitability (the ratio of current income to total assets), and liquidity (the ratio of liquid assets to total assets). In addition a vector of variables representing the main bank's lending capacity, *B_CAPACITY*, consisting of the bank's size (the natural logarithm of total assets), its financial health (risk-unadjusted capital/asset ratio), and profitability (the ratio of operating profits to total assets) is included. Finally, to control for industry-level shocks that affect firm investment, a set of industry dummies, *Industry*, is included.[9]

12.4.2 Export Behavior

To examine the impact of the loan supply shock resulting from the natural disaster on firms' export behavior, Miyakawa et al. (2014) adopt a methodology similar to the one employed by Minetti and Zhu (2011) and Koenig et al. (2010).

Specifically, let $Start_{it}$ denote a dummy that takes 1 if the firm starts exports in year t, and 0 otherwise, and $Export_{t-1}$ denotes the value of exports in year $t-1$. Then, assuming that the error term is normally distributed with zero mean and unit variance, the probability that firm i starts exporting conditional on that the firm did not export in year $t-1$ can be expressed as follows:

[9] Table 12.1 reports the results when firms are classified into six industries, namely, mining and construction; machinery manufacturing; other manufacturing; wholesale, retail and restaurants; finance, insurance, real estate, transportation, and communications; and other), and consequently five industry dummies are added. In the discussion paper version, firms are classified into five industries, yielding similar results.

$$\begin{aligned}\Pr\left(Start_t = 1 | Export_{t-1} = 0\right) = \Phi(&\beta_0 + \beta_1 F\_TFP_{it-1} + \beta_2 F\_DAMAGED_i \\ &+ \beta_3 B\_DAMAGED_i \\ &+ \beta_4 F\_DAMAGED_i * B\_DAMAGED_i \\ &+ \beta_5 F\_CONSTRAINTS_{i,t-1} \\ &+ \beta_6 B\_CAPACITY_{it-1} + \beta_7 Industry_i) \\ &for\ t = 1995, 1996, 1997. \end{aligned} \tag{12.2}$$

In this model, Φ denotes the standard normal cumulative distribution function (CDF). Equation (12.2) is estimated applying a linear probability model (i.e., OLS) to the sample of firms that did not export in year $t-1$.[10] The main interest of the analysis is in the effect of bank damage on the export decision of firms located outside the earthquake-hit area, which is captured by *B_DAMAGED*, because such bank damage is purely exogenous to these firms.

The dependent variable represents the probability of starting exports conditional on that the firm did not export in the previous period, as discussed above. The regressors include the firm's total factor productivity (TFP), following the pioneering theoretical work by Melitz (2003) and subsequent empirical studies (e.g., Bernard and Jensen 2004). The other explanatory variables are the same as in the investment Eq. (12.1) except for *F_SALESGROWTH*, which is omitted from Eq. (12.2).

12.4.3 Results

12.4.3.1 Capital Investment

The results for the baseline estimation for the investment ratio are shown in Table 12.1. For each year, the results for two specifications are reported: one using (1) *B_HQDAMAGED* and the other using (2) *B_BRDAMAGED* as the bank damage variable (referred to as *B_DAMAGED*). Note that only the variables of interest, *F_SALESGROWTH* and *B_DAMAGED*, are shown and all other control variables are omitted in Table 12.1. The results indicate that *F_SALESGROWTH*, the proxy for q, takes a positive coefficient in all years in both specifications, and is statistically significant in FY1995 and FY1997. Turning to the variables of primary interest, *B_DAMAGED* has a negative and significant coefficient in FY1995 (in specification (1)) and in FY1996 (in specification (2)), implying that the investment ratios of firms that were not hit by the earthquake were adversely affected if their main bank was

[10] The reason for not estimating (2) using probit or logit estimation is that the marginal effects and standard errors of the interaction term of *F_DAMAGED* and *B_DAMAGED* in these nonlinear models would need to be corrected (see, e.g., Norton et al. 2004). Angrist and Pischke (2008) have shown that the coefficients obtained using OLS estimation are virtually the same as the marginal effects using probit estimation.

Table 12.1 Year-by-year cross-section regressions for investment ratios

Dependent variable: *F_INVESTMENTRATIO* (t)	(1) *B_DAMAGED* = *B_HQDAMAGED*	(2) *B_DAMAGED* = *B_BRDAMAGED*	(1) *B_DAMAGED* = *B_HQDAMAGED*	(2) *B_DAMAGED* = *B_BRDAMAGED*	(1) *B_DAMAGED* = *B_HQDAMAGED*	(2) *B_DAMAGED* = *B_BRDAMAGED*
	FY1995		FY1996		FY1997	
F_SALESGROWTH (t-1)	0.0974[c] (0.0530)	0.1018[c] (0.0529)	0.0464 (0.0479)	0.0450 (0.0480)	0.1711[a] (0.0490)	0.1740[a] (0.0490)
B_DAMAGED[d]	− 0.0820[a] (0.0229)	− 0.0401 (0.0557)	− 0.0303 (0.0296)	− 0.1285[b] (0.0594)	0.1715[a] (0.0667)	0.0062 (0.0612)
Obs	1955	1955	1990	1990	1997	1997
F-value	8.88[a]	8.09[a]	6.65[a]	6.80[a]	8.48[a]	8.02[a]
R-squared	0.0812	0.0793	0.0468	0.0478	0.0581	0.0567
Root MSE	0.2223	0.2226	0.2239	0.2238	0.1997	0.1998
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes

Heteroskedasticity-robust standard errors are reported in parentheses. a, b, and c indicate significance at the 1 %, 5 %, and 10 % levels, respectively
[d]The *B_DAMAGED* variable is either *B_HQDAMAGED* or *B_BRDAMAGED* as indicated in the column heading

hit. As damage to banks is an exogenous financial shock for firms located outside the earthquake-hit area, this result strongly suggests that exogenous shocks to bank lending capacity affect client firm investment.

The impact of bank damage on undamaged firms is also economically significant. In specification (1), where bank damage is defined as headquarters damage, the investment ratio of undamaged firms whose main bank was damaged was 8.2 percentage points lower than that of undamaged firms whose main bank was not damaged. This impact is economically significant, given that the average investment ratio for undamaged firms in FY1995 was 13.1 %. In specification (2), where bank damage is represented by damage to bank branches, the investment ratios of undamaged firms with a damaged main bank, at the sample mean of *B_BRDAMAGED* for undamaged firms (i.e., 7.0 %), in FY1996 on average were 0.9 (= 7 × 0.1285) percentage points lower than those of firms with an undamaged main bank. Again, the quantitative impact is economically significant.

An interesting finding is that the timing of the impact of bank damage on firm investment differs between the two specifications. While the negative and significant impact of *B_HQDAMAGED* on client firms' investment manifested itself immediately after the earthquake, i.e., in FY1995, the significant impact of *B_BRDAMAGED* did so only one year later, in FY1996. This difference may stem from what these variables represent. *B_HQDAMAGED* captures the impairment of a bank's back-office operations at the headquarters, such as making decisions on whether to accept or reject applications for large loans, while *B_BRDAMAGED* reflects the damage to a bank's ability to process applications for small loans, and/or loan portfolio losses caused by the deterioration in local borrowers' financial conditions due to the earthquake. Note that the effects of bank damage, either to headquarters or to branch networks, are short-lived. The coefficient on *B_HQDAMAGED* turns *positive* and significant in FY1997, possibly reflecting a recovery or catch-up process from the low investment caused by bank damage in FY1995, while *B_BRDAMAGED* is not significant in FY1997.[11]

12.4.3.2 Exports

The results for the probability of starting exports conditional on firms not exporting in the previous period are shown in Table 12.2. Again, for each year the results for two specifications are reported: one using (1) *B_HQDAMAGED* and the other using (2) *B_BRDAMAGED* as the bank damage variable (referred to as *B_DAMAGED*). In addition to *B_DAMAGED*, only the results for *F_TFP* are shown in the table, while the results for all other control variables are dropped.

[11] The positive coefficient on *B_HQDAMAGED* in FY1997 may also reflect the survival bias arising from only selecting firms that survived over the three observation years. However, the dropout rate was not significantly different in FY1997 between firms with damaged main banks and firms with undamaged main banks.

Table 12.2 Year-by-year cross-section regressions for extensive margins

Dependent variable: *Prob*(*Start*(*t*) = *1*|*Export* (*t* − 1) = *0*) (*t*)	(1) *B_DAMAGED* = *B_HQDAMAGED*	(2) *B_DAMAGED* = *B_BRDAMAGED*	(1) *B_DAMAGED* = *B_HQDAMAGED*	(2) *B_DAMAGED* = *B_BRDAMAGED*	(1) *B_DAMAGED* = *B_HQDAMAGED*	(2) *B_DAMAGED* = *B_BRDAMAGED*
	FY1995		FY1996		FY1997	
F_TFP (t-1)	− 0.0328 (0.0312)	− 0.0324 (0.0313)	0.0030 (0.0247)	0.0044 (0.0246)	0.0191 (0.0202)	0.0190 (0.0200)
B_DAMAGED[c]	− 0.0436[a] (0.0139)	− 0.0480 (0.0522)	− 0.0510[a] (0.0143)	− 0.0844[b] (0.0484)	− 0.0237[a] (0.0075)	0.0297 (0.0401)
Obs *F*-value	1,993 6.29[a]	1,993 4.79[a]	1,917 4.60[a]	1,917 3.95[a]	1,953 3.02[a]	1,953 3.01[a]
R-squared Root MSE	0.0089 0.2097	0.0094 0.2096	0.008 0.1836	0.0088 0.1836	0.0036 0.1470	0.0037 0.1470
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes

Heteroskedasticity-robust standard errors are reported in parentheses. a and b indicate significance at the 1 % and 10 % levels, respectively
[c]The *B_DAMAGED* variable is either *B_HQDAMAGED* or *B_BRDAMAGED* as indicated in the column heading

The results indicate that *B_HQDAMAGED* has negative and significant coefficients for all three years following the earthquake. In contrast, *B_BRDAMAGED* does not take significant coefficients except for FY1996, for which it takes a negative and marginally significant coefficient. These results imply that the probability of starting exports for firms not directly hit by the earthquake was adversely affected if their main bank was hit. As damage to banks was an exogenous financial shock for firms located outside the earthquake-hit area, this result strongly suggests that exogenous shocks to bank lending capacity affected the probability of whether client firms started exporting.

As in the case of investment, the impact of bank damage on undamaged firms is also economically significant. In specification (1), where bank damage is defined as headquarters damage, the probability of starting exports for undamaged firms with a damaged main bank is 4.4 percentage points smaller than that of undamaged firms with an undamaged main bank in FY 1995. This impact is economically significant, given that the average probability of starting exports for undamaged firms in FY1995 was 4.4 %. The negative impact increases to 5.1 % in FY1996 and decreases to 2.4 % in FY1997. In specification (2), where bank damage is represented by damage to bank branches, the probability of starting exports for undamaged firms with a damaged main bank, at the sample mean of *B_BRDAMAGED* for undamaged firms (i.e., 6.8 %), in FY1996 had a 0.3 percentage point lower probability of exporting than firms with an undamaged main bank. Thus, the quantitative impact of damage to bank branch networks is not negligible, although it seems to be much smaller than the impact of damage to bank headquarters. It should also be noted that the impact of branch damage only appeared with a one-year lag.

12.5 Conclusion

This chapter presented an overview of the extant literature examining the causal link from fund supply shocks to various firm activities. To establish the existence of such a bank-lending channel, researchers need to isolate fund supply shocks from fund demand shocks. Recent research, including the two studies by the authors of this chapter presented here in detail, has attempted to overcome this important identification challenge by employing various identification strategies and datasets. Several studies have confirmed that shocks to loan supply have both statistically and economically significant impacts on firm activities such as investment and exports.

However, evidence based on successful identification strategies is still scarce and limited to a small number of economies and specific events such as financial crises or natural disasters. Bank lending affects firm activities when two financial frictions exist: (1) banks are unable to raise equity or obtain liquidity without significant costs when they suffer damage to their equity or liquidity; and (2) firms are unable to switch from damaged to healthier lenders without significant costs. Given that such financial frictions differ across economies and time, reflecting, for example, differences in the degree of financial development and financial stability, further evidence from different economies and events will help to provide a better understanding of the circumstances under which bank lending affects real economic activities. A closely related issue is what types of policies or institutions can mitigate the negative impact of a fund supply shock on firm activities. Finally, examining broader aspects of firm activities, including firm entry and exit, relocation, networking, and employment, will provide further guidance for a better understanding of the connection between bank lending and firm activities.

References

Amiti, M., & Weinstein, D. (2011). Exports and financial shocks. *Quarterly Journal of Economics, 126,* 1841–1877.

Amiti, M., & Weinstein, D. (2013). How much do bank shocks affect investment? Evidence from matched bank-firm loan data. *NBER Working Paper* 18890 (March 2013).

Angrist, J. D., & Pischke, J-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Ashcraft, A. B. (2005). Are banks really special? New evidence from the FDIC-induced failure of healthy banks. *American Economic Review, 95,* 1712–1730.

Bae, K-H., Kang, J-K., & Lim, C-W. (2002). The value of durable bank relationships: Evidence from Korean banking shocks. *Journal of Financial Economics, 64,* 181–214.

Bernanke, B. S. (1983). Nonmonetary effects of the financial crisis in the propagation of the great depression. *American Economic Review, 73,* 257–276.

Bernanke, B., & Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *American Economic Review, 79*(1), 14–31.

Bernanke, B. S., & Lown, C. S. (1991). The credit crunch. *Brookings Papers on Economic Activity, 2,* 205–48.

Bernard, A. B., & Jensen, J. B. (2004). Why some firms export. *Review of Economics and Statistics, 86,* 561–569.

Brewer III, E., Genay, H., Hunter, W. C., & Kaufman, G. G. (2003a). The value of banking relationships during a financial crisis: Evidence from failures of Japanese banks. *Journal of the Japanese and International Economies, 17,* 233–262.

Brewer III, E., Genay, H., Hunter, W. C. & Kaufman, G. G. (2003b). Does the Japanese stock market price bank-risk? Evidence from financial firm failures. *Journal of Money, Credit, and Banking, 35,* 507–543.

Caballero, R. J., Hoshi, T., & Kashyap, A. K. (2008). Zombie lending and depressed restructuring in Japan. *American Economic Review, 98,* 1943–1977.

Chava, S., & Purnanandam, A. (2011). The effect of banking crisis on bank-dependent borrowers. *Journal of Financial Economics, 99,* 116–135.

Chor, D., & Manova, K. (2012). Off the cliff and back: Credit conditions and international trade during the global financial crisis. *Journal of International Economics, 87,* 117–133.

Driscoll, J. C. (2004). Does bank lending affect output? Evidence from the U.S. states. *Journal of Monetary Economics, 51,* 451–471.

Fama, E. (1980). Banking in the theory of finance. *Journal of Monetary Economics, 6,* 39–57.

Fukuda, S., & Koibuchi, S. (2007). The impacts of "Shock Therapy" on large and small clients: Experiences from two large bank failures in Japan. *Pacific-Basin Finance Journal, 15,* 434–451.

Gibson, M. S. (1995). Can bank health affect investment? Evidence from Japan. *Journal of Business, 68,* 281–308.

Gibson, M. S. (1997). More evidence on the link between bank health and investment in Japan. *Journal of the Japanese and International Economies, 11,* 296–310.

Hancock, D., & Wilcox, J. A. (1998). The "Credit Crunch" and the availability of credit to small firms. *Journal of Banking and Finance, 22*(6–8), 983–1014.

Holmstrom, B., & Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics, 112*(3), 663–691.

Hori, M. (2005). Does bank liquidation affect client firm performance? Evidence from a bank failure in Japan. *Economics Letters, 88,* 415–420.

Hosono, K., & Miyakawa, D. (2014). Business cycles, monetary policy, and bank lending: Identifying the bank balance sheet channel with firm-bank match-level loan data. *RIETI Discussion Paper Series* 14-E-026.

Hosono, K., Miyakawa, D., Uchino, T., Hazama, M., Ono, A., Uchida, H., & Uesugi, I. (2012). Natural disasters, damage to banks, and firm investment. *RIETI Discussion Paper Series* 12-E-062.

Jiménez, G., Ongena, S., Peydro, J., & Saurina, J. (2012). Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *American Economic Review, 102,* 2301–2326.

Khwaja, A. I., & Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review, 98,* 1413–1412.

King, R., & Plosser, C. (1984). Money, credit, and prices in a real business cycle. *American Economic Review, 74,* 363–380.

Kiyotaki, N., & Moore J. (1997). Credit cycles. *Journal of Political Economy, 105*(2), 211–248.

Kline, M. W., Peek, J., & Rosengren, E. S. (2002). Troubled banks, impaired foreign direct investment: The role of relative access to credit. *American Economic Review, 92*(3), 664–682.

Koenig, P., Mayneris, F., & Poncet, S. (2010). Local export spillovers in France. *European Economic Review, 54,* 622–641.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71(6), 1695–1725.

Mian, A., & Sufi, A. (2011). House prices, home equity-based borrowing, and the U.S. household leverage crisis. *American Economic Review, 101,* 2108–2131.

Minamihashi, N. (2011). Credit crunch caused by bank failures and self-selection behavior in lending markets. *Journal of Money, Credit and Banking, 43,* 133–161.

Minetti, R., & Zhu, S. C. (2011). Credit constraints and firm export: Microeconomic evidence from Italy. *Journal of International Economics, 83,* 109–125.

Miyakawa, D., Hosono, K., Uchino, T., Ono, A., Uchida, H., & Uesugi, I. (2014). Financial shocks and firm exports: A natural experiment approach with a massive earthquake. *RIETI Discussion Paper* 14-E-010.

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *Stata Journal, 4*(2), 154–167.

Paravisini, D., Rappoport, V., Schnabl, P., & Wolfenzon, D. (2011). Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. *NBER Working Paper* 16975.

Peek, J., & Rosengren, E. S. (2000). Collateral damage: Effects of the Japanese bank crisis on real activity in the United States. *American Economic Review, 90,* 30–45.

Peek, J., & Rosengren, E. S. (2005). Unnatural selection: Perverse incentives and the misallocation of credit in Japan. *American Economic Review, 95,* 1144–1166.

Popov, A., & Udell, G. F. (2010). Cross-border banking and the international transmission of financial distress during the crisis of 2007–2008. *European Central Bank Working Papers* No. 1203.

Puri, M., Rocholl, J., & Steffen, S. (2011). Global retail lending in the aftermath of the US financial crisis: Distinguishing between supply and demand effects. *Journal of Financial Economics, 100,* 556–578.

Sawada, Y., & Shimizutani, S. (2008). How do people cope with natural disasters? Evidence from the Great Hanshin-Awaji (Kobe) earthquake. *Journal of Money, Credit, and Banking, 40,* 463–488.

Schnabl, P. (2012). The international transmission of bank liquidity shocks: Evidence from an emerging market. *Journal of Finance, 67,* 897–932.

Slovin, M. B., Sushka, M. E., & Polonchek, J. A. (1993). The value of bank durability: Borrowers as bank stakeholders. *Journal of Finance, 48,* 247–266.

Van Rijckeghem, C., & Weder, B. (2001). Sources of contagion: Is it finance or trade? *Journal of International Economics, 54,* 293–308.

Yamori, N. (1999). Stock market reaction to the bank liquidation in Japan: A case for the information effect hypothesis. *Journal of Financial Service Research, 15,* 57–68.

Yamori, N., & Murakami, A. (1999). Does bank relationship have an economic value? The effect of main bank failure on client firms. *Economics Letters, 65,* 115–120.

Kaoru Hosono is a professor of Economics at Gakushuin University. He is interested in macroeconomics and finance, with an emphasis in empirical research in these areas. His specific interests are in banking regulations, corporate finance, monetary policy, and aggregate productivity. He received his Ph.d. in Economics from Hitotsubashi University. He has published papers in the *Journal of the Japanese and International Economies, Japanese Economic Review, Japan and the World Economy,* and *Economics of Innovation and New Technology*, among others. His recent studies have been on the effects of natural disasters on firm dynamics including investment, export, entry/exit, and relocation, and on the transmission of monetary policy. He is a member of the Study group for Earthquake and Enterprise Dynamics (SEEDs).

Daisuke Miyakawa is an associate professor of Economics at Hitotsubashi University. His primary research is in financial institutions and firm dynamics. He received his Ph.d. in Economics from UCLA. He has published papers in the *Economic Inquiry, International Review of Finance,* and *Review of Monetary and Financial Studies*, among others. Using firm and financial institution match-level micro data, his recent studies involve the transmission of unconventional monetary policy, the impacts of the transaction partners' characteristics (e.g., banks and venture capital) on firm dynamics, and the matching mechanism of firms and financial institutions. He serves as a research member at Research Institute of Economy, Trade and Industry (RIETI), the Study group for Earthquake and Enterprise Dynamics (SEEDs), and the Canon Institute for Global Studies (CIGS).

Index

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks,* Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8