THE RELATIONSHIP BETWEEN COMPETITION AND INNOVATION: MEASURING INNOVATION AND CAUSALITY

ISBN 978-90-816238-0-3
Printed by Ridderprint, Ridderkerk

This book was typeset using LaTeX.

THE RELATIONSHIP BETWEEN COMPETITION AND INNOVATION: MEASURING INNOVATION AND CAUSALITY

De relatie tussen concurrentie en innovatie: het meten van innovatie en causaliteit
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus,
prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
dinsdag 24 mei 2011 des middags te 12.45 uur

door

Ryan van Lamoen
geboren op 21 oktober 1983 te Willemstad, Curacao

Promotor: Prof.dr. C.J.M. Kool

Co-promotor: Dr. J.W.B. Bos

Voor mijn ouders

Acknowledgements

Many people perceive the life as a PhD student as a long and lonely apprenticeship. In my experience, time seems to fly when you are having fun and I never felt lonely due to the frequent interaction with my supervisors, co-authors and other PhD students. I am very grateful to several colleagues, family members and friends and consider my PhD thesis the result of a team effort.

First, I thank my co-promotor Jaap Bos. It has been a real honor and pleasure to work with you on research projects. Your enthusiasm and innovative ideas have always inspired me to conduct research. I have learned a lot during our cooperation and your supervision and advice always guided me in the right direction. A PhD student cannot wish for a better supervisor since you were practically available for questions 24/7. This thesis substantially improved due to your advice and friendship.

Second, I thank my promotor Clemens Kool. When I was a student, Jaap once told a fellow-student and myself that he considers you to be one of the best economists in the Netherlands. After having you as a supervisor, I can say that I strongly agree with this statement. Our discussions, all the feedback on my research and your help with my personal development have made me a better economist.

I also take this opportunity to thank my co-authors of the research papers on which chapter 2, 3 and 5 are based. I thank Jim Kolari for his contributions on chapter 2 and all our informal conversations. It has been a real pleasure to work together with you thanks to your good sense of humor. I am also grateful to Claire Economidou for her help on chapter 3 and her friendship. Many thanks go to Mark Sanders for his contributions on chapter 5.

I thank the members of the reading committee, Prof. Dr. Rob Alessie, Prof. Dr. Jaap Bikker, Prof. Dr. Jan Boone, Prof. Dr. Pierre Mohnen and Prof. Dr. Arjen van Witteloostuijn for reading my thesis and their feedback. Your comments were very helpful for the final version of my dissertation.

Working on a PhD is not necessarily a lonely apprenticeship. It has also been a pleasant journey due to my roommates and fellow PhD students Joras Ferwerda, Lu Zhang, Yi Zhang and Secil Danakol.

I thank my best friend and fellow PhD student Martijn Dröes not only for all the work related discussions, but also for our good friendship. Your great sense of humor has definitely made my life as a PhD student more pleasant. Furthermore, our joint work has been very enjoyable and productive. If we become successful in the future, we may need to consider revealing the secret behind the success.

I thank my little brother Remi van Lamoen for all the fun when I visited him after work, for motivating me to stay in good shape after consuming too much junk food during work and for always being there in times of trouble.

I am forever indebted for the never-ending support of my parents, Richard and Karin van Lamoen. Since I was young you emphasized the importance of education, but provided me the freedom to follow my education interests. Your advice and love helped me in making all kinds of important decisions in my life and encouraged me to pursue my dreams.

Dear Jolyn, meeting you is one of the greatest blessings in my life. I am very grateful for your support during the last crucial year of my PhD track. Even though I did not always make it easy for you while finishing this thesis, you have always stood by me. I not only thank you for your support, but more importantly, I thank you for your love and the joy you bring into my life. If finding a passion for something or someone is the most important thing in life, I can proudly say that I have found it in you.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and background

Technological progress is an important engine of economic growth. Many studies find that firm performance is positively related to innovation activities such as R&D. Innovation activities may result in successful innovations which allow innovating firms to surpass their competitors and earn higher rents. While firms benefit from engaging in innovation activities, research also shows that the social return to R&D investments is higher than the private return. The social return to R&D may be the result of knowledge spillovers to other firms that are not involved in the actual creation of the knowledge. The recognition of the importance of technological progress for firm performance and economic growth has ignited the search for drivers of innovation. The search for these drivers and the seminal work of Schumpeter (1942) raised the interest in the relationship between competition and innovation among policymakers and researchers. In the Schumpeterian view, there is a positive relationship between monopoly power and innovation since high rents provide firms the advantage to finance innovations internally and the ability to hire the most innovative people (Kamien and Schwartz, 1982). Hence, competition is perceived to be detrimental for innovation activities since it not only reduces the monopoly rents that reward innovation, but also hampers advantages related to internal finance and the employment of innovative people. This negative effect effect of competition on innovation is also confirmed by empirical studies. However, several empirical studies also find that competition has a positive effect on innovation.[1] Aghion et al.

[1] See Symeonidis (1996) for an overview of empirical studies that find negative or positive effects of competition on innovation.

(2001) theorize that competition affects innovation positively since firms attempt to escape from their competitors by innovating if they are subject to more competitive pressures. Also inverted-U relationships have been found and suggest that there is an optimal level of competition that enhances innovation (Scherer, 1967). The empirical findings are mixed and depend on the data and empirical setting (industry or set of industries, linear or nonlinear specification).

The influential work of Aghion et al. (2005b) and Aghion and Griffith (2005) reconciles the mixed findings in the theoretical and empirical literature by developing a theoretical model that is able to explain an inverted-U relationship between competition and innovation.[2] Their model shows that competition affects innovation positively at lower levels of competition, where most firms are trying to escape from their competition. Competition influences innovation negatively at higher levels of competition since firms that lag behind the technological leaders are discouraged to innovate due to low post innovation rents.[3] Despite the developments in the competition-innovation literature over the past decades, many challenges remain for future research. These challenges primarily concern the measurement of the variables of interest, capturing causal effects in the competition-innovation relationship and dealing with (firm, industry) heterogeneity in the empirical setting.

Kamien and Schwartz (1982) argue that one of the most fundamental difficulties in the extant literature is the identification of an innovation.[4] Conventional measures such as patent statistics, innovation counts, R&D expenditures and R&D related personnel are often used in the existing literature, but have their limitations.[5] Furthermore, these innovation measures are not readily available for many sectors such as services and are mostly utilized in studies on the manufacturing sector. Relatively less attention

[2]The model incorporates a negative and a positive effect of competition on innovation to explain the existence of an inverted-U relationship. The model is also able to explain solely a positive or a negative effect of competition on innovation. This positive or negative effect depends on the relevant degree of competition in an industry.

[3]This corresponds with the argument from Schumpeter to the extent that more competition discourages innovation by reducing post-innovation rents. However, while all firms are subject to this mechanism in the model of Aghion et al. (2005b), only technological laggards experience a negative effect on their innovation behavior since firms with an equal level of technology will still try to surpass their competitors.

[4]While measuring competition is an important issue in this literature, the focus of this thesis is not on the measurement of competition.

[5]Not all innovations are patented, it is difficult to account for the economic value of patents and innovation counts and not all innovations require R&D expenditures and R&D personnel.

has been paid to innovation behavior in the service sector, while the service sector constitutes the most important part of the economic activities in developed countries. For example, the service sector in the U.S. accounted for 78.9% of U.S. GDP in 2002 and revenues from U.S. services relative to the rest of the world averaged 32.3% between 1980 and 2001 (Gallaher and Petrusa, 2006). Within the service sector, financial services comprise a large proportion of the economic activities. In 2005, finance, insurance and real estate accounted for 30.2% of the value added in the U.S. (Timmer and De Vries, 2009). Examining the influence of competition on firms' innovation behavior in these sectors is important given the crucial role of these sectors in the economy. Furthermore, innovations in financial services also foster an important engine of economic growth, namely financial development (King and Levine, 1993). Nevertheless, the effect of competition on innovation has yet to be explored in financial services and many other service sectors due to the scarcity of appropriate innovation measures. The existing studies in financial services only focus on specific technologies, do not focus on the competition-innovation relationship or do not account for firm heterogeneity in the innovation behavior of firms. For example, Hannan and McDowell (1984) focus on the relationship between market concentration and a specific technology, namely the adoption of ATMs. Hunter and Timme (1991) focus on a broader range of technological developments, but do not relate technological change to competition and do not allow firm firm heterogeneity in the innovation behavior of firms.[6] Another important drawback of these studies besides the measurement of financial innovation and/or accounting for firm heterogeneity is that they do not investigate causal effects.

Measuring innovation and accounting for firm/industry heterogeneity are not the only difficulties in the existing competition-innovation literature. There is a consensus that changes in market structure and competition affect the innovation behavior of firms. But how does innovation affect competition? For example, Microsoft has been profiling itself as an innovative player in the computer industry, but also raised concerns by U.S. antitrust authorities due to its dominance in the sector. This example illustrates that innovation may have an influence on competition since firms with successful innovations may surpass their competitors and become more dominant in the market. While most papers in the extant literature focus on the effect of competition on innovation, relatively less attention is paid to the effect

[6] These authors find that large banks experience more technological change than smaller banks, but assume a similar rate of technical change for all banks or groups of banks.

of innovation on competition (Symeonidis, 1996). The results are mixed since both negative and positive effects of innovation on competition have been found in the existing literature.[7] An important drawback of these studies is that they examine the effect of innovation input and output in isolation. Therefore, the empirical findings in the extant literature depend on the measurement of innovation since innovation input and output are interrelated and may convey different information on the degree of competition.[8] Therefore, examining the effect of innovation input and output in one framework and investigating the relationship between innovation output and input is important in the quest to capture the causal effect of innovation on competition. Examining this topic in detail is important from both a firm's perspective and a policy perspective. For example, theoretical and empirical contributions on this topic provide insights in the effectiveness of innovation behavior and performance of firms for their competitive position. Authorities in turn have an important tool at their disposal to reward innovative behavior of firms and to affect the degree of competition in sectors, namely patent laws.[9] Another tool of governments to affect the innovation and competition behavior of firms is the provision of R&D subsidies. These subsidies may foster innovation by inducing research that would not be conducted without these funds. Globalization and international competition increased countries' awareness of innovation as an instrument to gain a competitive edge.[10]

Despite the recognized importance of the relationship between competi-

[7]Symeonidis (1996) provides an overview of the literature.

[8]Innovation output affects competition if firms gain competitive advantages through successful innovations and innovation input may affect competition by altering firms' expectations of their competitive position in the future. Furthermore, there is a direct relationship between innovation input and output since innovations are produced by putting efforts in the creation of innovations.

[9]Intellectual property protection is an important determinant of the profit from an invention as it provides the inventor exclusive rights to exploit their inventions for a limited period of time. Successful innovators benefit from strict patent policies by protecting their innovations from imitation and through licensing. In countries with well-developed patent laws, patent infringements are often battled out in court. For example, Microsoft had to pay Alcatel-Lucent 367.4 million dollars due to violations of patents related to MP3-technology. Another example is the file suit by Nokia against Apple in 2009 due to violations of patents related to wireless technologies.

[10]For example, the European Union launched the Seventh Framework Programme for research and technological development (FP7) at the end of 2006 to fund research over the period 2007 to 2013 with a budget of 53.2 billion euros and a budget increase of 63% compared to their previous Framework Programme (FP6). The research programme is introduced by the European Union to increase the industrial competitiveness (Bruce et al., 2004).

tion and innovation activities by policymakers and researchers, the mixed findings in the competition-innovation literature suggest that potential channels and the innovation production process need to be examined further.

1.2 Aim of the thesis

This thesis focuses on the introduction of a new innovation measure and causal effects within the competition-innovation relationship. First, a new measure of innovation is introduced by examining the technology gaps between U.S. banks. The new innovation measure allows for analyses on the effect of competition on innovation in sectors where conventional innovation measures are not readily available.[11] Second, the thesis explores the competition-innovation relationship in banking more elaborately due to important market developments in the banking industry, such as the consolidation trend over the last decades. Specifically, the effect of scale economies on the competition-innovation nexus is examined in light of this consolidation trend and its impact on the market structure and competition. Third, the thesis explores the effect of innovation on competition based on Dutch innovation data. While competition affects the innovation incentives of firms, innovation activities and innovation output may in turn affect the degree of competition. The effect of innovation input and output on competition are disentangled in this thesis and examined in one framework to identify the relative contribution of these factors to competition. Examining these dimensions of innovation in one framework is necessary due to the relationship between innovation output and input. Fourth, this thesis focuses on examining the interrelatedness between innovation output and input closely by investigating the innovation production processes of Dutch firms and the role of efficiency in this production process.

1.3 Thesis outline

Chapter 2 introduces a new measure of innovation to examine the relationship between competition and innovation in the U.S. banking industry. The

[11] Frame and White (2004) argue that one of the main reasons for the lack of innovation studies in financial services is the unavailability of these conventional measures of innovation. One important objective of this thesis is to bridge the gap between financial economists and economists trained in industrial organization by using a new measure of innovation and to examine market developments and implications for competition in banking.

new innovation measure is obtained by estimating and enveloping annual cost frontiers to create a global frontier. The global frontier represents the best potentially available technology across banks and time. The distance to the global frontier represents the technology gap of each bank. Banks decrease their technology gap by implementing (process) innovations. The new measure of innovation is used to estimate the model of Aghion et al. (2005b) at the firm level for the U.S. banking industry in the period 1984-2004. This period is characterized by substantial consolidation and deregulation. Hence, the impact of the consolidation trend and interstate banking deregulation on competition and the innovation behavior of banks is examined. In addition, policy implications to financial reform and prudential regulation are discussed also.

To introduce a new measure of innovation, two strands of literature are combined. First, the thesis relies on the industrial organization and endogenous growth literature on competition and innovation.[12] Specifically, the analysis of the competition-innovation relationship in U.S. banking build on the influential work of Aghion et al. (2005b) and Aghion and Griffith (2005). Their model is applied to the banking industry to investigate the innovation behavior of banks and how they respond to changes in competition. Second, developments in efficiency analysis are used to introduce a new measure of innovation. Many studies on the estimation of production functions and efficiency use agriculture or manufacturing data and focus mostly on methodology.[13] The estimation of efficiency differences between banks also received the attention of economists due to these advances in econometric techniques and developments in the banking industry worldwide.[14] Recent developments in the broad efficiency literature by Battese et al. (2004), O'Donnell et al. (2008) and Bos and Schmiedel (2007) allow the estimation of technology gaps between firms. In general, innovations improve the technology set and affect technology gaps between firms, where successful innovators may surpass their competitors and become technological leaders and firms who fail to keep up with the technological pace lag behind. Technology gaps can be used to examine innovation in industries,

[12]Studies on the determinants of innovation are one of the largest fields in industrial organization and endogenous growth. Aghion and Griffith (2005) discuss early industrial organization models on the competition-innovation relationship and the endogenous growth paradigm in detail.

[13]Kumbhakar and Lovell (2000) provide an elaborate discussion of the literature and methodological developments.

[14]Berger et al. (1999) provide an overview of the efficiency literature in banking and developments in banking such as technological progress and consolidation trends.

where conventional measures such as patents and R&D are not readily available.[15] In contrast to these conventional innovation measure, the technology gap captures a broad range of innovations. Furthermore, while this innovation measure may capture other factors than innovations only, it has advantages over Total Factor Productivity since influences from efficiency and scale economies are controlled for.[16] Nevertheless, the new innovation measure in this thesis captures firm-specific innovation behavior and allows for a firm-level analysis on financial innovation. The existing firm-level studies on financial innovation are focused on specific technologies or do not allow for firm heterogeneity in the innovation behavior of firms.[17]

In chapter 3, the new innovation measure is utilized to examine the relationship between competition, scale economies and innovation. Specifically, the chapter is focused around the question whether large U.S. banks have become 'too big to innovate' . The theoretical work of Aghion et al. (2005b) is extended by relaxing their assumption that unit costs are independent from output levels in order to investigate the effect of scale (dis)economies on the competition-innovation nexus. The model is used to derive conditions under which the innovation behavior of firms with diseconomies of scale becomes more or less responsive to changes in competition. First, individual bank data over the period 1984-2004 are used to investigate the development of scale economies and competition in the consolidating U.S. banking industry. Second, the empirical analysis proceeds by investigating whether large U.S. banks have become 'too big to innovate' .

Chapter 4 is focused on the effect of innovation on competition. Specifically, the analysis is focused on whether innovation input and output have a different effect on competition. Most studies in the extant literature examine these dimensions of innovation in isolation and therefore cannot identify the relative importance of innovation input and output for the degree of competition.[18] The theoretical literature on the determinants of tacit collusion is used to examine the effect of innovation input and output on competition in one framework.[19] Specifically, as a basis for the empirical analysis on

[15] The estimation of technology gap requires production statistics or financial information.

[16] Aghion et al. (2005b) also calculate technology gaps based on Total Factor Productivity.

[17] For example, by assuming a similar rate of technical change for banks or groups of banks (Hunter and Timme (1991).

[18] The empirical evidence on this topic is mixed. Symeonidis (1996) provides an overview of the literature.

[19] Tirole (1988) and Ivaldi et al. (2003) provide an overview of the determinants of tacit collusion.

this topic, this thesis builds on the work of Tirole (1988) and Ivaldi et al. (2003) by extending their model to relate innovation input, the probability to innovate and innovation output to competition.[20] Ivaldi et al. (2003) only investigates the relationship between the probability to innovate and the sustainability of tacit collusion. Two-yearly Dutch firm-level data from the Community Innovation Survey and Production Statistics over the period 1994-2004 are utilized to investigate the effect of innovation input and output on competition empirically in one framework. This framework is also used to analyze whether innovation input or output is more important for competition. Furthermore, the analysis contains a comparison between the innovation-competition relationship in manufacturing and the service sector. Investigating the effect of innovation input and output on competition in one framework is not only necessary for the identification of the relative importance of the input and output, but also due to the relationship between innovation input and output.

Chapter 5 proceeds with a more detailed analysis of the relationship between innovation output and input. The aim of chapter 5 is to analyze the innovation production process and several sources of productivity and efficiency of this process. In an influential paper by Mairesse and Mohnen (2002), they estimate an innovation production function and emphasize the importance of innovativeness, which is the unexplained ability to turn innovation inputs into innovation output. However, their analyses do not focus on the role of efficiency as a component of innovativeness. Based on two-yearly data from the Dutch Community Innovation Survey over the period 1994-2004, stochastic frontier analysis is used to estimate innovation production functions and to identify inefficiency. In addition, the productivity of the innovation inputs and inefficiency are linked to cooperation with competitors, cooperation with other institutions, funding from the government, competition and firm size. Different strands of literature are combined to investigate the sources of productivity and efficiency in the innovation production process. In particular, the literature on knowledge production functions is combined with the literature on efficiency analyses. Studies on the innovation production function and inefficiency are scarce and mostly performed at the macro-level (Wang, 2007; Wang and Huang, 2007). As in this thesis, Gantumur and Stephan (2010) also analyze the production of innovation and inefficiency at the firm-level, but use different estimation techniques and do not focus on the same research questions (e.g., they do not examine the effect of government funding).

[20] These relationships are not examined in Tirole (1988) and Ivaldi et al. (2003).

Summing up, a new measure of innovation is introduced in chapter 2 to examine the relationship between competition and innovation in the U.S. banking industry, where conventional innovation measures are not readily available. In chapter 3, the theoretical and empirical framework of Aghion et al. (2005b) is extended and the new innovation measure is used to investigate how the presence of scale (dis)economies affect the competition-innovation relationship in U.S. banking. The effect of innovation input and output on competition is analyzed in chapter 4. In chapter 5, the relationship between innovation output and inputs is examined in more detail by estimating innovation production functions. Finally, chapter 6 contains a conclusion based on the evidence found in the thesis. Furthermore, limitations of the analyses and directions for future research are discussed.

Chapter 2

Competition and Innovation: Evidence from Financial Services

2.1 Introduction

Seminal work by Schumpeter (1942) posits that product market competition discourages innovation by diminishing monopoly rents. By contrast, Aghion, Harris, Howitt, and Vickers (2001) assert that competition may foster innovation as firms attempt to escape competition.[1] Providing partial support for both conjectures, some empirical studies find an inverted-U pattern between competition and innovation (e.g., Scherer, 1967; Levin and Mowrey, 1985). In an attempt to reconcile theory and evidence, Aghion and Griffith (2005) and Aghion, Bloom, Blundell, Griffith, and Howitt (2005b) propose a theoretical model that is able to explain an inverted-U relationship between competition and innovation, wherein an escape competition effect initially dominates until competition reaches a sufficient level at which the rent dissipation effect thereafter prevails.[2]

The purpose of this chapter is to examine the relationship between competition and innovation by extending the previous literature from manufacturing to financial services. We contribute to the competition/innovation literature in two major ways. First, we introduce a new overall measure of financial innovation. Instead of using traditional innovation outputs (e.g.,

[1]See literature reviews by Kamien and Schwartz (1982) and Symeonidis (1996).

[2]Their empirical evidence for manufacturing industries in the U.K. tends to support the hypothesis of an inverted-U pattern.

patents which are mostly relevant to manufacturing), we examine banks' ability to minimize costs through innovations. Following earlier work by Hayami and Ruttan (1970), Mundlak and Hellinghausen (1982), and Lau and Yotopoulos (1989), we estimate and envelope annual minimum cost frontiers to create a global frontier. The distance to the global frontier constitutes each bank's technology gap, which decreases if the bank manages to innovate. The use of this innovation measure enables us to examine the innovation behavior of financial firms for which patents and R&D expenditures are unavailable as metrics. Each innovation leads to lower production costs in both the theoretical model and our proposed innovation measure. Because a well-functioning financial sector is crucial to the economy (King and Levine, 1993; Pagano, 1993; Levine, 1997; Levine and Zervos, 1998; Levine et al., 2000; Levine, 2004), it is important to understand its innovation dynamics. While numerous studies have been published on bank innovation, they typically focus on specific bank technologies, rather than bank innovation in general. For example, Hannan and McDowell (1984) examine how market concentration affects the adoption of ATMs. Our technology gap measure provides estimates of overall innovation to gain a broader perspective on its potential effects. Also, while most competition/innovation studies employ industry-level data and therefore implicitly assume the same relation across industries despite considerable cross-industry heterogeneity, our new measure affords the opportunity to study competition/innovation behavior at the firm level for a single industry.

Second, we document the impact of historic consolidation in the U.S. banking industry on innovation behavior. Driven by globalization, technological change, deregulation, and other forces, the U.S. banking industry has experienced dramatic changes in its structure and competition (Jones and Critchfield, 2005). In our sample period from 1984 to 2004, the industry consolidated from over 14,000 banks to around 7,500 banks (Berger, Kashyap, and Scalise, 1995; Berger, Demsetz, and Strahan, 1999). In this period, the average size of banks grew as banks with assets totaling more than $10 billion increased their share of industry assets from 30 percent to over 70 percent (Rhoades, 2000).[3] A major concern is that consolidation has lowered competition. For example, Cetorelli and Strahan (2006) find that, in banking markets that are highly concentrated, nonfinancial firms have significantly less access to credit (see also Stiroh and Strahan, 2003). Also, Fraser et al. (2011) report evidence of market power after large bank

[3]Similar consolidation trends have occurred in the European Union, United Kingdom, Japan, and other countries around the world (Carletti et al., 2007).

mergers that adversely affects the stock prices of borrowing firms. Relevant to our purpose, a number of interesting questions naturally arise. As consolidation has taken place, what are the coincident trends in U.S. bank competition and innovation over time? How has consolidation affected the relationship between competition and innovation? And, what is the role of deregulation in bank innovation dynamics? Our empirical analyses provide detailed evidence on these important questions.

In brief, based on annual data series for all insured U.S. commercial banks spanning two decades, our evidence strongly supports an inverted-U relationship between bank competition and technology gaps. This finding is consistent with the theory and evidence by Aghion et al. (2005b) and Hashmi (2007) in other industries. We also find that average price cost margins have increased considerably during our sample period, which implies declining competition as consolidation has occurred. Further evidence suggests that: (1) the U.S. banking industry as a whole has consolidated beyond its optimal innovation level, and (2) interstate banking deregulation has lowered innovation through its effect on competition. In view of these adverse trends, we discuss potential implications to policy makers making sweeping financial reforms and changes in prudential regulations at the present time.

Section 2.2 provides a brief overview of studies on financial innovation in the U.S. banking industry. Section 2.3 describes the theoretical model developed by Aghion et al. (2005b) to explain the inverted-U pattern. Section 2.4 overviews the data and methodology. Section 2.5 empirically investigates the existence of an inverted-U relationship, discusses model robustness, considers whether the consolidation process has gone too far, and examines the impact of interstate banking deregulation. Section 2.6 concludes.

2.2 Financial innovation in U.S. banking

Deregulation of prices, products, and geographic restrictions on permissible banking activities over the past 30 years has increased the contribution of market forces to financial innovation in the banking industry. In this regard, Miller (1986) argues that efforts to circumvent regulatory and tax burdens are key drivers of financial innovation. Also, Vives (2001) observes that deregulation and financial innovations, including advances in information technology, management techniques, and risk adjustment (e.g., derivatives, securitization, and off-balance sheet activities), have substantially increased competition in U.S. and European banking markets.

Frame and White (2004) comprehensively survey the small body of financial innovation literature comprised of 39 empirical studies. They define financial innovation as comprising activities that internally reduce bank costs and risks or externally better meet the convenience and needs of customers.[4] Financial innovations are grouped into new products (e.g., automated teller machines or ATMs, credit and debit cards, adjustable-rate mortgages, etc.), new production processes (e.g., electronic payments and record keeping, automated credit scoring models, securitization of loans, etc.) and new organizational forms (e.g., interstate banking organizations, diversified banks with traditional and nontraditional financial services, etc.). The practical significance of these financial innovations lies in their contribution to enhancing financial intermediation, which allocates savings to investment and thereby contributes to economic growth (see King and Levine (1993), Levine (1997), and others cited above). Frame and White (2004) conclude that the following factors tend to increase innovation in financial services: regulation, institution size, higher education and income, and first-mover, cost, and reputational advantages. Given the important role of financial innovation in the financial system and the economy as a whole, they infer that there is considerable room for future research in this "relatively untilled field."

A separate branch of the banking literature relevant to this study examines technical change in the context of cost and profit efficiency analyses of financial institutions.[5] In general, the efficiency literature tends to support the institution size effect in financial innovation cited by Frame and White (2004). Elyasiani and Mehdian (1990) and Hunter and Timme (1991) find that larger banks experienced greater cost efficiency gains compared to small banks in the 1980s. Humphrey (1993) finds that large banks had more technical change than small banks in the late 1970s. Also, Berger and Mester (1997) find that, while large banks had decreasing cost efficiency over time, they exhibited increasing profit efficiency compared to small banks in the 1980s and 1990s.[6] Consistent with these studies, Wheelock and Wilson

[4]Van Horne (1985) more broadly defines financial innovation as making markets more operationally efficient or complete (i.e., the number and types of securities that span all possible return and risk contingencies or states of the world demanded by market participants). Also, Allen and Gale (1994) propose that financial innovation is associated with efficient risk sharing due to the completion of markets.

[5]For example, Van Horne (1985) observes that financial innovations are motivated by operational inefficiencies. Less efficient financial institutions are less competitive and, therefore, less likely to survive. Importantly, as Ross (1989) points out, institutions are the major agents of innovation in financial markets.

[6]Similarly, Berger and Mester (2003) report decreasing cost productivity but increasing profit productivity among U.S. banks in the period 1991-1997.

(1999) report greater technological gains among large banks in the 1980s and 1990s, which led them to conclude that competitive and regulatory changes in the banking industry have benefited larger over smaller banks. Lastly, Altunbas, Goddard, and Molyneux (1999) find that larger banks in 15 European countries experienced more gains from technical change than smaller banks in the period 1989-1996.

While the above studies find that institution size is positively related to financial innovation in line with Schumpeter's hypothesis, few empirical studies have attempted to link financial innovation and competition. Based on a sample of about 3,800 U.S. commercial banks in the period 1971-1979, Hannan and McDowell (1984) report evidence that the likelihood of ATM adoption is positively related to bank size, market concentration (i.e., three-firm concentration ratios in SMSAs or counties), and membership in a bank holding company. They infer that Schumpeter's hypothesis is supported by these empirical results. Chourchane, Nickerson, and Sullivan (2002) employ bivariate and multivariate logit analyses to examine the competitive effects of internet banking for about 1,600 U.S. commercial banks in 1999. Unlike Hannan and McDowell, they find that the market concentration of competitive rivals lowers the likelihood of banks entering internet banking markets. Also, faced with uncertain demand, larger banks are more likely to enter internet banking than smaller banks, who prefer to delay their investment decision until larger banks have committed assets to the technological change. Another study by Mantel and McHugh (2001) considers the question of whether, given regulatory oversight of consumer protection issues, there is sufficient competition and innovation in consumer electronic payments, including credit cards, debit cards, e-cash, and smart cards. They conclude that private sector efforts to achieve adequate consumer safety are as effective as regulatory intervention. Also, financial innovation in electronic consumer payments would increase to a greater extent due to market forces than regulation. Finally, Akhavein, Frame, and White (2005) test for the impact of market competition on the adoption of small business credit scoring for a sample of 96 large U.S. banks in 1997. Like prior studies, bank organization size is positively related to early adoption of credit scoring methods. However, market concentration as measured by an average Herfindahl-Hirschman index (HHI) across local geographic markets is not significant. Overall, these studies yield mixed evidence on the link between competition and innovation but confirm larger banks are innovation leaders.

Other studies posit a variety of theories concerning innovation and financial services. Van Horne (1985) conjectures that uncertainty about regulations, tax laws, inflation, international events, and technology will lead to a continuing stream of financial innovations. Boot and Thakor (1997) theorize that a universal banking system comprised of joint commercial and investment banks will produce less innovation than functionally-separated financial institutions due to adverse spillover effects of (for example) securities innovations on commercial banking profits. However, they argue that in mixed financial systems with both universal and functionally-separate institutions, large universal banks will have a competitive advantage to influence changes in regulations that favor financial innovations in which scope economies are required.[7] Related work by Bhattacharyya and Nanda (2000) theorizes that larger investment banks will be more likely to innovate new financial services due to larger market shares with greater revenue incentives. Smaller banks have less incentive to innovate but are expected to be more aggressive than large banks in their introduction strategy (e.g., attracting large bank customers). And, another study by Hauswald and Marquez (2003) proposes that financial innovation in information processing and dissemination can change competition in the banking industry (see also Wilhelm, 2001). Advances in information technology that improve information processing capabilities of banks will tend to decrease competition as some banks gain an information advantage over other banks. Conversely, if information dissemination is increased by information technology innovations, competition will increase due to widespread access to proprietary information that levels the playing field. They conclude that the relative importance of these two information effects is an empirical question.

In sum, consistent with empirical evidence, theoretical studies predict a variety of competition and innovation relationships for financial services, in addition to a greater likelihood of financial innovation by larger institutions. Previous empirical studies typically focus on a particular innovation, rather than total innovation by financial institutions. The main reason for this limitation is the inability to measure overall innovation. We seek to contribute to the financial innovation literature by proposing and estimating a general measure of innovation referred to as the *technology gap*. Our new innovation measure enables tests of theories about competition and innovation as well as new insights into the innovativeness of the U.S. banking industry.

[7] The stakeholder capture theory proposes that some agents can sway public policy decisions in their favor (e.g., see Kroszner and Strahan, 1999).

2.3 The Model

This section applies the model developed by Aghion et al. (2005b) to the banking industry. The model is used to derive the "escape competition effect" (positive effect of competition on innovation) and "Schumpeterian effect" (negative effect of competition on innovation), which together explain an inverted-U relationship between competition and innovation. Aghion et al. model an economy and derive the average flow of innovations for intermediate sectors. We explain and apply their model to the banking industry and derive the average flow of innovations for a bank operating in one or several geographical banking markets to provide more coherency between the theoretical model developed by Aghion et al. (2005b) and the empirical analysis in this chapter.[8]

The Aghion et al. model assumes that there is a continuum (with total mass equal to one) of identical consumers in the economy that use a constant intertemporal discount rate r and have the utility function $u(y_t) = \ln y_t$. Here we extend their model to the banking industry, where the final good y_t (financial services) is produced using input services from a continuum of intermediate sectors according to the production function $\ln y_t = \int_0^1 \ln x_{jt} dj$, where x_{jt} is an aggregate of two intermediate goods A and B produced by two banks (duopoly) in sector j. Each intermediate sector represents a geographical (e.g., local) banking market.[9] The subutility function is defined as $x_j = x_{Aj} + x_{Bj}$. Consumers maximize the subutility function with respect to their (normalized) budget constraint.[10]

It is assumed that the bank production function exhibits constant returns, and banks use labor as an input at the (exogenous) normalized wage rate $w(t) = 1$. The unit cost of production is independent of the quantity produced, such that the unit cost structure becomes $c_i = \gamma^{-k_i}$, where γ^{-k_i} is the unit labor requirement of bank i, γ is the size of an innovation (assumed to be larger than one), and k_i is the technological level of a bank. Hence, innovations lower the unit cost by decreasing the required units of labor per unit of output. The relative costs of a bank depend only on the technolog-

[8] The theoretical and empirical analysis of Aghion et al. (2005b) are conducted at the industry level, while this chapter focuses on competition and innovation behavior at the firm level.

[9] Competition in banking occurs at a local level for many products and services (e.g., see Pilloff, 1999, and Berger et al., 1999). For example, the model in this paper may represent the whole U.S. banking industry as an aggregate of the financial services in local markets.

[10] The income of consumers is normalized to unity by using expenditure as the numeraire for prices in each period.

ical gap. The maximum technological gap in a sector is assumed to be one ($m = 1$), and technological advances occur through step-by-step innovations instead of leapfrogging models.[11] The R&D cost function $\psi(n) = n^2/2$ is expressed in units of labor n. Furthermore, a Poisson process for innovations is assumed and laggards or neck-and-neck banks move one technological step ahead with a Poisson hazard rate of n (R&D intensity) by spending $\psi(n)$ on R&D. The laggard bank moves ahead with the hazard rate $n + h$ if it puts effort into R&D, where h is a help factor that represents R&D spillovers or the ability to copy the technology of a leader. The R&D intensities of leading banks, laggard banks, and neck-and-neck banks are n_1, n_{-1}, and n_0, respectively. By assumption, leaders do not innovate ($n_1 = 0$), as laggard banks can copy their previous technology immediately (so that the maximum gap remains one).

Product market competition is modeled by the ability of banks to collude. It is assumed that banks are able to collude if they are operating in a leveled bank market segment but cannot collude if the bank market segment is unleveled. It is important to note that multi-market contact may facilitate collusive behavior.[12] We assume that a bank may face another bank in several markets, but it does not face the same bank in all markets. Furthermore, it is assumed that a bank facing one competitor in several markets has the same technology gap in these markets.[13] The profits of laggard and leader banks are $\pi_{-1} = 0$ and $\pi_1 = 1 - \gamma^{-1}$, respectively. Laggard banks make zero profit, as leaders capture the market and earn a profit equal to their revenue (normalized to one due to the income in the budget constraint) minus the costs (equal to the inverse of the innovation parameter γ). The profit of neck-and-neck banks ranges from zero to one-half of the profits of a technological leader. The inability to collude leads to zero profits, as banks are assumed to be in Bertrand competition with undifferentiated products and similar unit costs, $\pi_0 = \varepsilon\pi_1, \varepsilon \in [0, \frac{1}{2}]$. Competition is parameterized by $\Delta = 1 - \varepsilon$ and equals the incremental profit of an innovating bank in a

[11] It is impossible for laggard banks to surpass a technological leader by means of an innovation without drawing even with this leader. See Aghion et al. (1997) for several appealing features of a model of step-by-step innovation compared to the Schumpeterian leapfrogging models.

[12] For example, it is possible to give a competitor a higher market share in one market to induce collusive behavior, while using the other market for disciplining purposes.

[13] Suppose that banks A and B are competitors in two local markets. If bank A is a technological leader in one market, this bank can use its technological advantage in both markets. This assumption implies that markets are either leveled or unleveled. Our treatment here is similar to what Park and Pennacchi (2009) use for large multi-market banking organizations.

leveled market normalized by the profit of a leader.

Aghion et al. derive the research intensities and examine how they are affected by changes in competition. It is assumed that the discount rate is zero ($r = 0$). The research intensities of neck-and-neck banks and laggards are $n_0 = \sqrt{h^2 + 2\Delta\pi_1} - h$ and $n_{-1} = \sqrt{h^2 + n_0^2 + 2\pi_1} - h - n_0$, respectively.[14] Differentiating these research intensities with respect to the competition parameter Δ gives (see Aghion et al. 2005b, p. 722):

$$\frac{\partial n_0}{\partial \Delta} = \frac{\pi_1}{n_0 + h} > 0. \tag{2.1}$$

$$\frac{\partial n_{-1}}{\partial \Delta} = \frac{\partial n_0}{\partial \Delta}\left[-1 + \frac{n_0}{n_{-1} + h + n_0}\right] < 0. \tag{2.2}$$

Obviously, from the first order condition in equation (2.1), the research intensity of a neck-and-neck firm is positively affected by increases in the degree of competition (i.e., an escape-competition effect). Their innovation incentives increase with more competition, as their pre-innovation rents are reduced more than post-innovation rents. The first order condition in equation (2.2) shows that the research intensity of a laggard bank decreases as competition increases (i.e., a Schumpeterian effect). The reason for this negative effect on innovation is that more competition reduces the rents that a laggard bank can attain by innovating. Whether the escape competition or Schumpeterian effect dominates depends on the fraction of leveled and unleveled sectors in the steady state, as determined by the research intensities of laggards and neck-and-neck banks.

We diverge from Aghion et al. by examining how the average flow of innovations of a firm (as opposed to industry) changes with competition. The steady-state probabilities that a market is leveled and unleveled are μ_0 and μ_1, respectively. During any unit time interval, the steady-state probability that a market changes state (from leveled to unleveled or vice versa) is an aggregate of the probability of being a certain type of market times the Poisson hazard rate that firms move ahead. For unleveled and leveled markets this probability is $\mu_1(n_{-1} + h)$ and $2\mu_0 n_0$, respectively. The condition $\mu_1(n_{-1} + h) = 2\mu_0 n_0$ must hold in the steady-state, as the fraction of leveled and unleveled markets must remain unchanged. The probability of being a laggard bank (branch) or leader in a certain geographical market

[14] See Aghion et al. (2005b) for a derivation of equilibrium research intensities as well as escape competition and Schumpeterian effects.

is $p\mu_1$ and $(1-p)\mu_1$, respectively.[15] The average flow of innovations for a bank or branch in a geographical market is:[16]

$$I = \mu_0 n_0 + p\mu_1 \left(n_{-1} + h\right) + \left(1-p\right)\mu_1 n_1 = \frac{\left(1+p\right) 2n_0 \left(n_{-1} + h\right)}{2n_0 + n_{-1} + h}, \quad (2.3)$$

which equals the sum of the probability of being a certain type of bank (branch) times the research intensity.[17] This average flow of innovations in a certain geographical market follows an inverted-U pattern.[18] Moreover, the average flow of innovations from the perspective of bank I_T that operates in S geographical markets is the sum of the average flow of innovations in each geographical market j:

$$I_T = \sum_{j=1}^{S} I_j. \quad (2.4)$$

The intuition behind the inverted-U relationship is straightforward. If the degree of competition is initially low, neck-and-neck banks earn high profits and have little incentive to innovate. By contrast, laggard banks have relatively more incentive to innovate due to low initial profits (i.e., zero in the model) but high potential profits if they manage to catch up with a technological leader. We can infer that banks will leave their status as a neck-and-neck (laggard) bank relatively slowly (rapidly). Consequently, a bank will be a neck-and-neck bank most of the time, such that the escape-competition effect dominates. If there is not much competition in a market, increased competition should lead to a higher average innovation rate. The reverse is true in the case of high initial competition. Now laggard banks do not have much incentive to innovate due to little gain after a successful innovation. However, neck-and-neck banks have relatively more incentive to innovate due to large incremental potential profit. These outcomes imply that banks will leave their status as a neck-and-neck (laggard) bank relatively rapidly (slowly). In this scenario, a bank will be a laggard bank most of the time, such that the Schumpeterian effect dominates and the leader never innovates. Hence, if the degree of competition is initially high, increased competition should lead to a lower average innovation rate.

[15] The probability that a bank or branch operates in an unleveled sector $p\mu_1 + (1-p)\mu_1 = \mu_1$.

[16] Since two neck-and-neck firms are trying to gain a technological lead, the average flow of innovations in an intermediate sector is equal to $I = 2\mu_0 n_0 + \mu_1 \left(n_{-1} + h\right)$.

[17] The term $\left(1-p\right)\mu_1 n_1 = 0$, since it is assumed that leaders do not innovate ($n_1 = 0$).

[18] See Aghion et al. (2005b) for the proof at the industry level.

2.4 Data and Methodology

2.4.1 Data

Year-end 1984-2004 data are gathered for individual U.S. insured commercial banks from Call Reports of Income and Condition provided by the Federal Reserve System. Data are expressed in 1984 U.S. dollars.

Table 2.1: Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.
Price cost margin	0.179	0.090	-0.993	0.964
Total assets (millions of USD)	458.740	7416.978	1.067	967,365
Risk (Equity/Total Assets)	0.096	0.034	7.36e-05	0.998
Salary expenses per fte (thousands of USD)	35.143	12.756	0.048	537.160

The descriptive statistics are based on the sample of the preferred specification in Table 2.3 (specification 2). The number of observations is 151,476.

Table 2.1 shows the descriptive statistics for banks' price cost margin, total assets, equity ratio, and average wage per fte-employee.[19] The number of banks declined from about 14,000 in 1984 to 7,500 in 2004.[20]

2.4.2 Estimating technology gaps

Our approach to measuring innovation via shifts in the cost function (.i.e., technical change) has considerable precedent in previous empirical literature (Subramanian and Nilakanta, 1996; Ruttan, 1997; Agrell et al., 2002; Bleaney and Wakelin, 2002). Also, our approach closely aligns with the theoretical concept of technology gap posited by Aghion et al. (2001) and Aghion et al. (2005b).

Metrics for measuring technical change have evolved over time. Led by Tinbergen (1942), the econometric approach utilizes a time trend when estimating the cost (or production) function. Likewise, the index number theory approach of Solow (1957) identifies neutral technical change with constant marginal rates of substitution. As shown in Figure 2.1a, these approaches capture innovation by means of a parallel shift in the cost curve, in which other parameters of the cost function are unchanged.

[19]The total number of observations is approximately 220,000. We obtain 151,476 observations in the preferred specification due to missing values of certain variables, the exclusion of outliers, applying first-differences and using lags of the endogenous regressors as instruments.

[20]We assume that attrition of the panel data occurs exogenously.

Diewert (1976) and others extend the index numbers approach and add flexibility to the measurement of technical change by relaxing the assumption that the latter was constant. Likewise, Baltagi and Griffin (1988) extend the econometric approach and introduce a general index of technical change that allows for nonconstant technical change. Building on advances by Gollop and Jorgenson (1980) and others, they further allow for biased technical change, as marginal rates of substitution are allowed to vary over time. Consistent with these approaches, Figure 2.1a shows a nonparallel shift in the cost curve is possible.[21]

Since the cost function describes the *process* whereby firms are assumed to minimize costs in producing output, the set of estimated parameters reflects the state of technology. What separates technical change from other means of minimizing costs (such as adjusting the input mix or reducing waste) is the fact that, whereas the latter measures use the currently implemented technology, technical change consists of the invention or adoption of new technology. In an attempt to reconcile this view of technical change with the notion of estimating cost functions, early work by Hayami and Ruttan (1970), Mundlak and Hellinghausen (1982) and Lau and Yotopoulos (1989) introduce a meta frontier approach. The meta frontier (global frontier) encompasses the set of available technologies across firms and/or across time. Technical change consists of the adoption of a new technology as measured against the benchmark meta frontier, which combines all available technologies.

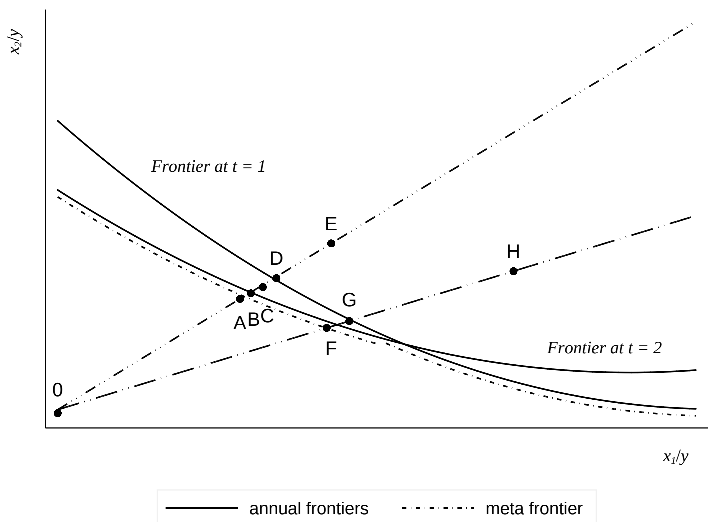
Figure 2.1a illustrates the notion of technical change using a simple paradigm with two inputs (x_1, x_2) and one output (y) for two firms I and II. In this example, there are two annual frontiers corresponding to $t = 1$ and time $t = 2$. Each frontier represents the minimum cost curve based on available technology for a certain level of output for that year. The cost efficiency of firm I located at point E at time $t = 1$ is OD/OE. If firm I is at point C at $t = 2$, its efficiency is OB/OC. Figure 2.1a also shows that firm II is located at point H and faces a cost efficiency of OG/OH at time $t = 1$. The dashed line that envelops the annual frontier represents the minimum cost frontier over the whole period, or meta frontier. Innovation results in a lower gap between the annual minimum cost frontier and the meta frontier.

Innovation is reflected by changes in the technology gap, which measures the difference between currently available technology and optimal technology

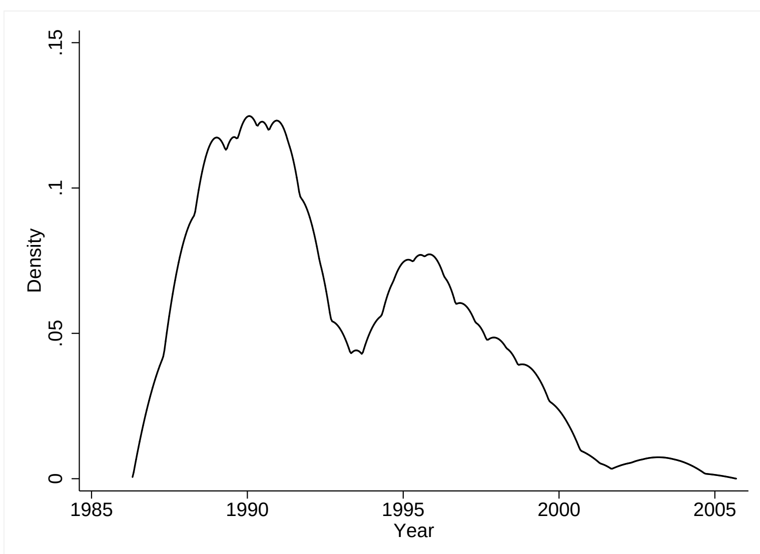
[21] The relationship between biased technical change and innovation is explained in Acemoglu (2002).

Figure 2.1: Technology gap ratio

(a) Meta frontier theory



(b) Distribution of banks with a technology gap of 1



over the whole period with values between zero and one (i.e., the firm is on the meta frontier).[22] At $t = 1$ the firm faces a technology gap of OA/OD, which narrows to OA/OB at $t = 2$ as the firm improves its technology set. While firm I faces a technology gap of OA/OD in period $t = 1$, the technology gap of firm II is smaller (OF/OG).

We follow recent work by Battese et al. (2004), O'Donnell et al. (2008) and Bos and Schmiedel (2007) and obtain technology gaps by initially employing Stochastic Frontier Analysis (SFA) to estimate the minimum cost frontier available in each year and then enveloping the annual cost frontiers to obtain a meta frontier (see also Hayami and Ruttan (1970), Mundlak and Hellinghausen (1982) and Lau and Yotopoulos (1989)).[23]

In the first step, the following annual translog cost frontiers are estimated using stochastic frontier analysis:[24]

$$TC_{it} = f^*(w_{it}, y_{it}, z_{it})_{it} e^{v_{it}+u_{it}}, \qquad (2.5)$$

where w represents the vector of input prices, y is the output vector, z is a vector of control variables, v is random noise assumed to be i.i.d. $N(0, \sigma_{vit})$, and u is the inefficiency term assumed to be i.i.d. $N|(\mu_{it}, \sigma_{uit})|$. To take into account inefficiency, we use stochastic frontier analysis, which is ignored by conventional measures of productivity (e.g., TFP) that measure technical change as efficiency change.[25]

We assume that banks minimize total costs and operate in perfectly competitive input markets. Bank production is modeled using the well-known intermediation approach.[26] As such, output y consists of loans, investments, and off-balance sheet items.[27] Input prices w correspond to the prices of fixed assets, labor, and borrowed funds.[28] The equity ratio z is included as a control variable to account for different risk profiles of banks.[29] The com-

[22]The notion of technology gaps was first used by Krugman (1979) and proxied in the literature by total factor productivity (TFP) differentials (Griffith et al., 2004).

[23]Kumbhakar and Lovell (2000) provide an elaborate discussion of the development and application of SFA to efficiency measurement.

[24]Homogeneity of degree one in input prices and symmetry are imposed.

[25]See Bos et al. (2009) for a discussion.

[26]See Freixas and Rochet (1997) for a more elaborate discussion on measuring the activities of banks.

[27]Output quantities are year-end stocks.

[28]Input prices are computed as follows: the price of fixed assets is depreciation divided by fixed assets, the price of labor is personnel expenses divided by the number of fte-employees and the price of borrowed funds is interest expenses divided by total borrowed funds.

[29]See Hughes and Mester (1993).

posed error term in equation (2.5) is $\varepsilon_{it} = \nu_{it} + u_{it}$. Firm-specific inefficiency estimates u are obtained by using the expected value of u_{it} conditional on the total error ε_{it} (i.e., $E(u_{it}|\varepsilon_{it})$). Cost efficiency score estimates are obtained as follows:

$$C\hat{E}_{it} = [\exp(-\hat{u}_{it})], \tag{2.6}$$

where CE equals one for banks that operate on the annual frontier (no inefficiency). Banks with inefficiencies operate above the annual cost frontier and have cost efficiency scores less than one.

In the second step, the meta frontier is estimated as the envelope around the annual cost frontiers. We utilize the parameter estimates for the annual cost frontiers and obtain estimates of the technology gap (GAP) by fitting the minimum cost meta frontier (f_{meta}) as follows:

$$\begin{gathered} Min.\text{Distance}_{it} = \sum_{t=1}^{T} \sum_{i=1}^{N} |lnf^{*}(w_{it}, y_{it}, z_{it})_{it} - lnf_{meta}(w_{it}, y_{it}, z_{it})_{it}| \\ s.t.\ lnf_{meta}(\cdot)_{it} \leq lnf^{*}(\cdot)_{it}. \end{gathered} \tag{2.7}$$

In this constrained minimization problem, the absolute distance between the annual cost frontier and the meta frontier is minimized subject to the constraint that the total cost from the annual frontier is equal to or larger than total cost from the meta frontier. As a result, the technology gap is defined as:

$$GAP_{it} = \frac{f_{meta}(w_{it}, y_{it}, z_{it})_{it}}{f^{*}(w_{it}, y_{it}, z_{it})_{it}}. \tag{2.8}$$

Innovations by firms may lead to improvements in their technology set and, consequently, a smaller gap between the current technology set and the (potentially available) best technology set, or meta frontier. The result is an increase in GAP_{it}, which is bounded between 0 and 1, where the latter is reached when firms operate on the meta frontier. Table 2.2 shows the descriptive statistics for the estimated technology gaps.

We can now relate the role innovation plays in the model of Aghion et al. (2005b) to the technology gap derived in equation (2.8) and depicted in Figure 2.1a. Recall their model crucially assumes that leaders do not innovate and that the maximum gap remains one. Additionally, innovations are expected to lower the unit cost of production (process innovations),

Table 2.2: Technology gap ratios

Variable	Mean	Std. Dev.	Min.	Max.
Technology gap	0.989	0.029	1.07e-08	1.000

The descriptive statistics are based on the sample of the preferred specification in Table 2.3 (specification 2). The number of observations is 151,476.

laggards cannot surpass the leader without first drawing even, and neck-and-neck firms may innovate and move the technology frontier forward. Clearly, the technology gap derived above satisfies these conditions.

The technology gap of a leader that is positioned on the global frontier in two consecutive periods will maintain a technology gap equal to one. Likewise, a laggard can close the technology gap by lowering his cost (catching up) and moving towards the global frontier (lowering the technology gap). Neck-and-neck firms may operate on the annual cost frontier under the best potential available technology in the current period and, subsequently, shift the annual cost frontier towards the global frontier in the next period by improving their technology set through innovations. In sum, the technology gap as a measure of overall innovation agrees with with the concept of innovation proposed by Aghion et al. (2005b).

Figure 2.1b shows the distribution of banks over the period 1984-2004 with a technology gap equal to one. These banks were operating on the global frontier and thus utilized the best potential technology set over the period given their combination of outputs and input prices. Most banks had a technology gap equal to 1 during the late 1980s and early 1990s compared to the other years. By contrast, there were relatively fewer banks operating at the best available technology in the early to mid-2000s. It should be noted that some banks may innovate in a certain year, while other banks introduce similar technologies or adopt the same technology in later years. This behavior may result in higher technology gaps for banks that innovate earlier and an increase in the technology gap for banks that innovate (or adopt the same technology) in later periods.[30]

2.4.3 Measuring competition

We follow Aghion et al. (2005b) by using the price cost margin (viz., Lerner index or markup) as the main indicator of competition and subtract it from one as follows:

[30]For example, Hannan and McDowell (1984) find that large banks have a higher conditional probability of adopting ATMs.

$$C_{it} = 1 - \left(\frac{\Pi_{it} + F_{it}}{R_{it}}\right), \tag{2.9}$$

where Π_{it} is the profit of a bank, F_{it} represents fixed costs, and R_{it} denotes sales. The price cost margin is calculated by dividing the net income after taxes and extraordinary items plus expenses of premises and fixed assets by total non-interest income plus total interest income. We prefer to use a firm-specific measure of competition instead of a measure based on a certain geographical market, as some banks compete mainly at the local level and others mainly at the national or international level. Hence, we assume that all changes in competition are reflected in the price cost margins of banks regardless of the geographical market in which banks are competing.[31] The competition measure ranges between zero and two after removal of outliers.[32]

2.4.4 Empirical specifications and estimation procedures

We consider several empirical specifications based on the following model:[33]

$$GAP_{it} = \beta_1 C_{it} + \beta_2 C_{it}^2 + \gamma' \mathbf{Z}_{it} + a_i + \varepsilon_{it}, \tag{2.10}$$

where the technology gap is the dependent variable, C_{it} is the competition variable, γ' is a 1xn parameter vector, and $\mathbf{Z}_{it}$ is a nx1 vector of control variables. A squared term for the competition variable is included to account for the inverted-U relationship between competition and innovation proposed by Aghion et al.. Taking the first-differences of equation (2.10) to eliminate the unobserved heterogeneity a_i gives:

[31] An important assumption in this thesis is that less competition is positively related with higher price cost margins, while more competition leads to lower price cost margins.

[32] Outliers were removed after visual inspection of scatter plots of the technology gap against the price cost margin. A range between -1 and 1 for the price cost margin was considered to be reasonable. Therefore, the analysis excludes 1,324 observations (i.e., less than 1% of the total amount of observations). Some authors choose to remove negative price cost margins, but this approach creates a bias in the results as only firms with positive price cost margins are considered. The dataset contains 14,073 observations with negative price cost margins. Excluding negative price cost margins also results in an inverted-u relationship, but the domain where the Schumpeterian effect dominates is smaller. Our empirical findings were robust to different thresholds.

[33] While we focus on process innovations by calculating technology gaps based on cost frontiers, the same framework can be used to examine the effect of competition on product innovations. Process innovations affect the incremental profit from innovation by lowering costs and product innovations by increasing revenues. Competition affects these profits and hence the innovation incentives of a firm.

$$\Delta GAP_{it} = \beta_1 \Delta C_{it} + \beta_2 \Delta C_{it}^2 + \gamma' \Delta \mathbf{Z}_{it} + \Delta \varepsilon_{it}. \tag{2.11}$$

The competition variable may be endogenous due to reverse causality with the innovation variable.[34] To deal with the endogeneity of the competition variable, it is necessary to find relevant instruments (i.e., correlated with the endogenous variable) that are not correlated with the error term (instrument exogeneity). We use the two-step efficient generalized method of moments estimator (GMM), where lags of the endogenous variables in levels are used as instruments for the endogenous variables in first-differences.[35] The lag structure will depend on the order of serial correlation in the residuals. If there is no serial correlation in the residuals (in levels), lags from period t-2 (and onwards) can be used as instruments. However, if there is first-order serial correlation (in the residual in levels), lags from period t-3 (and onwards) can be used. An important assumption is that these lags are not correlated with the disturbance term.[36]

Additionally, we employ a model specification in which interstate banking deregulation at the state level is used as an instrument. Before the late 1970s, restrictions on interstate banking protected banks from outside competition. Deregulation concerning interstate banking unleashed competitive forces by allowing banks to enter new markets and pose a threat to incumbent banks (Stiroh and Strahan, 2003). This regime shift is a dummy variable that takes the value one from the year in which states entered into an interstate banking agreement with other states, and zero before this year.[37] However, it is questionable whether such policy reforms in the banking sector are suitable instruments for competition. Kroszner and Strahan (1999) argue that several technological and financial innovations influenced

[34]Chapter 4 examines the effect of innovation output and innovation input on competition.

[35]OLS estimations were performed for exploratory purposes, but the OLS estimator gives biased and inconsistent estimates of causal effects in the presence of endogenous regressors. GMM has some efficiency gains compared to the traditional IV/2SLS estimator. For example, the two-step GMM estimator utilizes an optimal weighting matrix that minimizes the asymptotic variance of the estimator. Also, GMM is more efficient than the 2SLS estimator in the presence of heteroskedasticity.

[36]An argument against the exogeneity of these lags is that the dependent variable in the current period may reflect expectations in the past. If this is the case and behavior in the past is based on expectations of the future, these lags as instruments are not exogenous. For example, the technology gap ratio in the current period may reflect expectations concerning the technology gap in previous periods, and this expectation may have affected competitive behavior in the past.

[37]See Stiroh and Strahan (2003) for an overview of the deregulation year for each state.

the deregulation process by affecting the lobby behavior of banks. In this regard, these authors cite some specific technologies (e.g., the introduction of the ATM) that spurred banks' efforts to seek deregulation. Even though the technology gap ratio captures a broad spectrum of implemented innovations, the endogeneity of this regime shift due to reverse causality with technological developments remains questionable.

One empirical drawback of neglecting control variables in equation (2.10) is that other factors that are correlated with competition can influence innovation. Therefore, we introduce a model specification with several control variables that allow for the possibility of nonzero elements in the parameter vector γ. Our control variables are equity divided by total assets, firm size in terms of total assets, and the average wage per fte-employee. Equity to total assets is an inverse measure of debt pressure. Aghion et al. (2005a) argue that debt pressure is positively related to innovation, as firms increase innovation to escape the threat of bankruptcy. The relationship between firm size and innovation is a Schumpeterian hypothesis. Plausible explanations of a positive relationship between firm size and innovation are potential scale economies in R&D as well as diversification benefits that lower risk (Kamien and Schwartz, 1982).[38] Lastly, the variable average wage per full-time equivalent employee proxies for labor productivity and, therefore, is positively related to innovation since it is more likely that innovations are produced by more productive employees.

We also estimate equation (2.11) for two different time periods to examine the stability of the inverted-U pattern over time. The Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 eliminated interstate banking restrictions at the national level and is used to demarcate two periods: 1984-1993 and 1994-2004.

2.5 Results

This section describes the empirical results. First, we examine whether an inverted-U relationship exists. Second, we investigate the robustness of our results by comparing them to alternative specifications.[39] Third, and last, we evaluate the effect of the consolidation process on innovation in U.S.

[38] In earlier works, Schumpeter argued that individual entrepreneurs (small firms) were mostly responsible for innovations. Later on, he argued that large firms provide a conducive environment for innovations.

[39] We also performed a Kernel regression to allow for more flexibility in the relationship between competition and the technology gap ratio. The Kernel regression also shows evidence of an inverted-U relationship between competition and innovation.

banking and investigate how the interstate banking deregulation process, which was aimed in part at enhancing competition, affected innovation.

2.5.1 Is there an inverted-U relationship?

Our main focus is to examine the effect of competition on innovation in U.S. banking. Table 2.3 shows the results for the preferred and alternative specifications (denoted 1 and 2-6, respectively). The results in Table 2.3 suggest a statistically significant inverted-U relationship between competition and technology gap ratios in all model specifications. The competition variable and its squared term are individually and jointly significant at the 1% level. The preferred model specification 1 shows the results with lagged values in levels of the competition variable and its squared term.[40] Since the Arellano-Bond test for serial correlation indicates first-order autocorrelation in the residuals in levels, the lags are taken from periods t-3 and t-4.[41] According to the Hansen test, the null hypothesis that the instruments are valid cannot be rejected.[42] The optimal price cost margin in preferred specification 1 is around 5.2%. Translated into 1984 dollars, a one percentage point decrease in the price cost margin results on average in approximately \$627,000,000 lower costs for the whole banking sector (evaluated at the average number of banks per year in the sample of about 11,000).[43]

---

[40] An important drawback of the two-step efficient GMM estimator is that the standard errors are downward biased in small samples. Windmeijer (2005) proposes a finite sample corrected estimate of the variance. The corrected variance leads to more accurate inference in small samples. There is also a finite sample bias of the two-step GMM estimator itself. Therefore, the correction of the variance is only useful for improving inference when the estimator does not contain a large finite sample bias. However, these issues may not be a problem with the dataset used in this study due to the large number of observations. As a robustness check, we performed regressions with the Windmeijer correction. As expected, the correction had a negligible effect on the standard errors.

[41] This serial correlation test is based on an examination of residuals in first differences. Testing for first-order serial correlation in levels is based on testing for second-order serial correlation in first differences.

[42] We also performed an instrument relevance test based on the joint significance of the instruments. The F-statistics for the regressions with the competition variable and its squared term as a dependent variable are 905.16 and 545.13, respectively. Since an F-statistic of 10 is often used as a rule of thumb to examine instrument relevance, we conclude that the instruments are relevant.

[43] This estimate is obtained by evaluating the marginal effect of the competition variable on the technology gap ratio at the average price cost margin. The marginal effect on the technology gap ratio is translated into annual frontier dollar values evaluated at the average global frontier, average technology gap ratio, and the average technology gap ratio plus the marginal effect.

Table 2.3: Competition and technology gaps

Specification	1 *(preferred)*	2	3	4	5	6
Estimation procedure	*2-step GMM*	*OLS*	*2-step GMM*	*2-step GMM*	*2-step GMM*	*2-step GMM*
Period	1984-2004	1984-2004	1984-2004	1984-2004	1984-1993	1994-2004
$\Delta Competition_{it}$	1.600***	0.0600***	1.474***	1.424***	0.678***	2.511***
	(0.095)	(0.009)	(0.106)	(0.069)	(0.070)	(0.180)
$\Delta Competition^2_{it}$	-0.844***	-0.038***	-0.783***	-0.753***	-0.336***	-1.369***
	(0.058)	(0.005)	(0.062)	(0.042)	(0.041)	(0.121)
$\Delta Equity_{it}/total\ assets_{it}$			-0.124***			
			(0.037)			
$\Delta Total\ assets_{it}(\$1,000,000)$			-8.51e-08			
			(6.40e-08)			
$\Delta Average\ wage\ per\ fte_{it}$			-5.88e-05			
			(5.74e-05)			
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.000	0.000	0.000	0.000	0.002	0.000
AR(3)	0.969	0.100	0.795	0.814	0.798	0.899
Optimal price cost margin	5.2%	21.8%	5.8%	5.4%	-0.01%	8.3%
Hansen J statistic	1.064		0.797	2.089	1.151	1.071
	(0.588)		(0.671)	(0.148)	(0.562)	(0.585)
Observations	151,476	198,785	151,476	166,437	65,020	86,456

Based on equation (2.11). Standard errors (in parentheses) are robust against heteroskedasticity and serial correlation. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01. The p-values are reported for Arellano-Bond serial correlation tests. Chi-squared statistics and associated p-values are reported for the Hansen test.

Figure 2.2a shows the empirical relationship between the technology gap ratio and competition measure based on the preferred model specification 1. These findings are consistent with the theoretical and empirical results of Aghion et al. (2005b) and Hashmi (2007). Aghion et al. use U.K. data (industry averages) over the period 1973-1994 and find an optimal price cost margin around 6.0% after using policy instruments to deal with the endogeneity of competition (see Aghion et al. 2005b, Table 1, specification 4, p. 708).[44]

To examine the robustness of our results, we check whether an inverse-U relationship exists for a number of alternative model specifications. Table 2.3 gives the results. As shown there, model specification 2 shows the OLS estimation results without dealing with the endogeneity of the competition variable. The optimal price cost margin is around 21.8%. However, the OLS estimator yields biased and inconsistent estimates of the causal effect of the regressor on an outcome in the presence of endogenous regressors. The competition variable and its squared term are endogenous from a theoretical point of view due to reverse causality with innovation.[45]

Model specification 3 includes several control variables, and lagged values of the competition variable (from periods t-3 and t-4) are used as instruments. An inverted-U relationship between competition and technology gap ratio is again obtained after controlling for other factors that may affect innovation. According to the Hansen test, the instruments are valid in this specification. The optimal price cost margin is around 5.8% and thus higher than the optimal markup from the preferred specification. The risk variable's (equity divided by total assets) estimated coefficient is negative and significant at the 1% level. As such, a decrease in the equity ratio is associated with a higher technology gap ratio. This finding is consistent with Aghion et al. (2005a), who propose that debt pressure may lead to more innovation. Both firm size and the average wage per fte-employee are not significantly related to the technology gap ratio.

In model specification 4, the lag structure of the instruments is changed

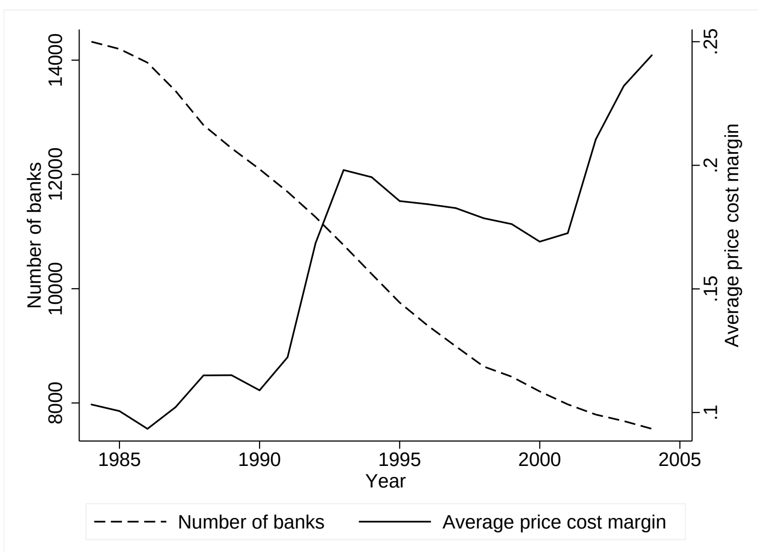
[44] Hashmi (2007) also conducts an industry-level analysis based on data from publicly-traded manufacturing firms in the U.S. over the period 1970-1994. Instead of using an instrumental variable approach, he uses the first lag of the competition variable and the squared term directly. His empirical results indicate an optimal price cost margin around 22.9% (see Hashmi, 2007, Table 2, specification 1, p.13).

[45] We also performed the Hausman-Wu endogeneity test based on preferred model specification 1 to examine the endogeneity of the competition variable and its squared term. The endogeneity test showed that the competition variable and its squared term should be treated as endogenous regressors.

Figure 2.2: The competition-innovation relationship

(a) The inverted-U relationship

(b) The average price cost margin and the number of banks



by using lags of the competition variable and its squared term from period t-3. A regime shift associated with interstate banking deregulation is added to the instrument set, such that there is one over-identifying restriction. The instruments are valid based on the Sargan-Hansen test.[46] We find a statistically significant (individually and jointly) inverted-U relationship between competition and the technology gap ratio, with an optimal markup of approximately 5.4%. Model specifications 5 and 6 show the results for time periods before and after the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994, respectively. Again, the inverted-U pattern is robust in these two time periods.

2.5.2 How have the consolidation trend and interstate banking deregulation affected competition and innovation in U.S. banking?

In an oral statement before the U.S. House of Representatives, Ludwig (1997) cites the Riegle-Neal Act as a successful example of how removing restrictions promotes competition and states that this "... increased competition should benefit consumers and businesses through lower costs, increased access, improved services and greater innovation." More recently, Zarutskie (2006) has investigated the effect of the Riegle-Neal Act on firm borrowing and investment. She concludes that banking market reforms such as Riegle Neal have a differential impact, which depends on the effects that the reforms have on (process) innovations in the industry, as the latter play an important role in bank loans to more opaque and younger firms. Extending our earlier results from specifications 4-6 concerning an inverted-U pattern for competition and innovation, we next consider whether the consolidation trend and interstate banking deregulation was associated with a change in competition and innovation (i.e., a shift along the horizontal axis in Figure 2.2a).

Comparing the results for specifications 5 (for the period 1984-1993) and 6 (for the period 1994-2004), we observe that the optimal price cost margin increased significantly, from -0.01% to 8.3%. Figure 2.2b shows the average price cost margin and number of banks over the sample period 1984-2004. In this period the number of banks declined from 14,323 in 1984 to

[46] According to the Arellano-Bond serial correlation test, there is no evidence of first order serial correlation in the residuals in levels. Hence, it would be appropriate to use the lag from period t-2 as an instrument also. We have chosen lags from period t-3, as many specifications showed some evidence of first-order serial correlation and therefore contradicted the results from this test in specification 4.

7,548 in 2004, while the average price cost margin increased from 9.3% in 1984 to 24.5% in 2004. The average price cost margin in 1984 was already higher than the optimal markup in our preferred specification (or 5.2%) and most of our other model specifications (except for the OLS results). However, the average price cost margin declined in some years in our sample period. For example, the average price cost margin declined every year from 1993 at 19.8% to 2000 at 16.9%. This means that movements along the inverted-U relationship were toward the optimal point that enhances innovation. Nevertheless, the average price cost margins soared again in 2001 to 2004 from around 17.3% to approximately 24.5%. This translates into a movement away from the optimal point, given that banks were positioned on average to the left side of the inverted-U relationship during the whole sample period. Thus, we find that the consolidation process in the U.S. banking industry has been accompanied by: (1) a large increase in average price cost margins, and (2) movement away from the optimal point that enhances innovation.

How has interstate banking deregulation affected competition and innovation? In model specification 4 the interstate banking deregulation was used as an instrument for competition.[47] This dummy variable assumes that the effect of deregulation was absorbed immediately in the competitive environment. To allow for flexibility in the response function, we follow Stiroh and Strahan (2003) and construct a dummy variable for every two consecutive years after the (state-specific) deregulation year in which the state entered into an interstate banking agreement with other states. The response function is estimated by regressing the competition variable on the full set of dummy variables that captures the dynamic effects of deregulation:[48]

$$C_{it} = \sum_{k=1}^{7} \alpha_k D_{kst} + a_i + \varepsilon_{it}, \tag{2.12}$$

where C_{it} is the competition variable, and D_{kst} is the deregulation variable that has been in effect for period k in state s. Each period consists of two consecutive years starting from year 1 until year 13. Only the last period consists of more than two years and captures the effect of deregulation from

[47] We also estimated a specification in which we regressed the technology gap on the competition variables and the deregulation variable as a control variable. The deregulation variable was insignificant in this specification.

[48] The equation is estimated in first-differences to eliminate the unobserved heterogeneity that is constant over time.

year 13 and thereafter in the post-implementation period. Table 2.4 and Figure 2.3 show the results.

Table 2.4: Competition and the dynamic effect of interstate banking deregulation

Estimation procedure	*OLS*
$\Delta Years\ \ 1-2$	-0.000
	(0.001)
$\Delta Years\ \ 3-4$	-0.003**
	(0.001)
$\Delta Years\ \ 5-6$	-0.007***
	(0.002)
$\Delta Years\ \ 7-8$	-0.024***
	(0.002)
$\Delta Years\ \ 9-10$	-0.026***
	(0.002)
$\Delta Years\ \ 11-12$	-0.029***
	(0.002)
$\Delta Years\ \ 13$ *thereafter*	-0.037***
	(0.002)
Observations	202,168

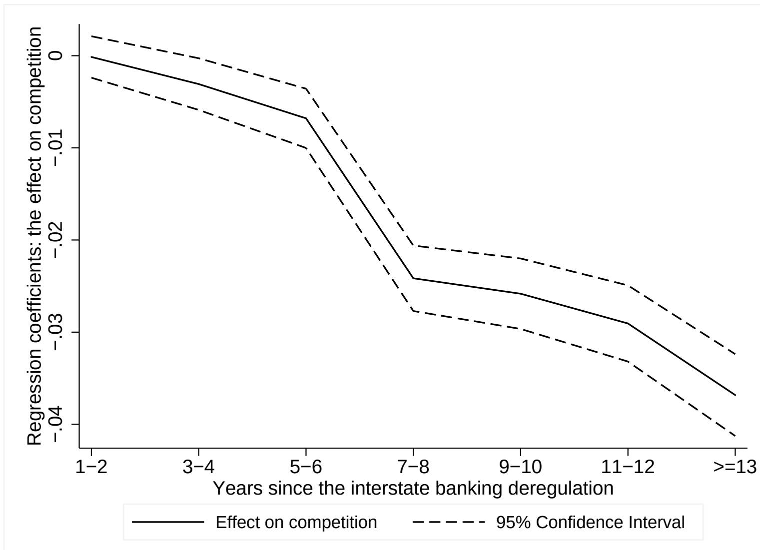
These results are based on equation (2.12). Standard errors (in parentheses) are robust with respect to heteroskedasticity and serial correlation. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01.

Most regression coefficients are negative and suggest that deregulation of interstate banking had an adverse effect on competition.[49] We performed an F-test to investigate whether the responses are homogenous over time. The null hypothesis that the slope coefficients are equal to each other is rejected at the 1% level.[50] A possible explanation for the decrease in competition is that more multi-market contact (due to interstate banking deregulation) facilitated tacit collusion. Our results are consistent with the findings of Stiroh and Strahan (2003), who argue that banking deregulation triggered a reallocation of banking assets from low profit to high profit banks resulting in increased profitability in the banking sector. Hence, the reallocation process in the industry led to higher price cost margins on average. The coefficients in Table 2.4 show a clear downward trend in the response of competition to

[49] There seems to be no significant effect in the first two years after deregulation.

[50] The p-value of the F-test is zero.

Figure 2.3: Competition and the dynamic effect of interstate banking deregulation



interstate banking deregulation. This evidence suggests that the negative competitive effects of the regime shift increased over time.

To make inferences about its effects on innovation, we need to know the level of competition before deregulation. The average price cost margin before the year of deregulation is 10.5%. This is higher than the optimal price cost margin in all of the specifications, except for specification 2 in which endogeneity problems are not treated. In view of the inverted-U relationship between competition and innovation, we infer that the technology gap ratio decreased on average after the deregulation of interstate banking (i.e., the leftward movement begins left of the optimal point). In sum, the interstate banking deregulation process appears to have led to a reduction in innovation in the U.S. banking industry.

2.6 Conclusion

This chapter seeks to contribute to the financial innovation literature by examining the relationship between competition and innovation in the U.S. banking industry. We develop and estimate a new measure of overall innovation by estimating technology gap ratios obtained from global frontier analyses of U.S banks. Year-end data for all insured U.S. commercial banks in the period 1984-2004 are employed. Consistent with Aghion et al. (2005b) and others, our results suggest an inverted-U relationship between competition and innovation in the U.S. banking industry. This relationship is robust to alternative model specifications using control variables and different instruments, in addition to different sample periods. Further analyses indicate that interstate banking deregulation negatively affected competition on average. The downward trend in competition after this deregulation implies a leftward movement along the inverted-U curve. Since banks were below (or left of) the optimal level of competition in our sample period, interstate banking deregulation tended to reduce innovation on average. Moreover, our results show that the consolidation movement has been accompanied by marked increases in the price cost margins of banks over time (e.g., from 9.3% in 1984 to 24.5% in 2004). These margins are higher than the optimal price cost margins in most of our model specifications.

At the present time, sweeping financial reforms and changes in prudential regulatory policies are being made by governments and regulatory agencies around the world. During the recent global financial crisis, massive government intervention in many countries sought to stabilize the financial system via mergers of large failing banks and other financial institutions. Given this consolidation of large, troubled banks, as well as the average position of U.S. banks on the inverted-U curve, we infer that further diminution of competition in the U.S. banking industry could negatively affect financial innovation in the years to come. Our results point to a a need to consider the systematic effects on the financial system of reforms and regulatory changes, such as how competition and innovation are affected on the industry level, rather than focusing on individual banks. Regulatory policies promoting competition would tend to reverse declines in innovation, reduce the size of very large institutions whose failure might threaten national and global economies (i.e., too-big-to-fail firms, or TBTFs), and foster safety and soundness of not only individual institutions but the financial system as a whole. Future research on the competitive and innovation effects of recent crises on banking institutions is recommended to corroborate adverse trends identified in this chapter and analyze policy options to improve industry conditions.

Chapter 3

Too Big to Innovate? Scale (dis)economies and the Competition-Innovation Relationship in U.S. Banking

3.1 Introduction

During the last two decades, the banking sector has changed profoundly. Globalization, advances in information technology, mergers and acquisitions and consolidation have reshaped the banking sector as the number of banks in the U.S. declined by nearly fifty percent (Berger et al., 1995, 1999). At the same time, the average size of banks has increased substantially, as banks with assets totaling more than ten billion dollars increased their share of banking industry assets from thirty percent to over seventy percent (Rhoades, 2000). Similar consolidation trends have occurred in the European Union, Japan, and other countries (Carletti et al., 2007).

Increased consolidation raises concerns, as more concentrated markets are generally believed to facilitate collusion (Berger et al., 2004).[1] But

[1] The relationship between market structure and competition has attracted considerable attention in the literature. However, the results are mixed. For example, Berger and Hannan (1989) find a robust positive relationship between profitability and market concentration in retail banking markets in the late 1980s. In contrast, Cole et al. (2004) report no evidence that differences in loan approval procedures of large banks versus small banks had a negative effect on pricing and volume in the market for small business lending. In tests of how competition in local banking markets affects the market structure of

consolidation may have other negative consequences as well. If large banks operate with increasing average costs, this affects their innovation incentives, as future profits from innovations are dampened by a further increase in average costs. In turn, a less innovative banking sector may raise the costs of financial intermediation and have negative effects on economic growth.[2]

Despite the widespread recognition of the rapid and broad proliferation of financial innovation and the relative abundance of innovation studies for other sectors of the economy (manufacturing, agriculture), there is a relative dearth of empirical studies on financial innovation.[3] Even less is known about the relationship between competition and innovation in banking.

Nonetheless, the competition-innovation nexus itself has recently received a boost by some important advances. Whereas the seminal work by Schumpeter (1942) posits that competition discourages innovation by diminishing monopoly rents that result from innovation (the so-called 'Schumpeterian effect'), recent work by Aghion et al. (2001) argues that competition may foster innovation as firms attempt to escape competition (the so-called 'escape competition effect') by engaging in innovative activities.[4] In an attempt to reconcile these theories and mixed empirical evidence, Aghion et al. (2005b) have proposed a theoretical model that establishes an inverted-U relationship between competition and innovation, wherein an escape competition effect initially dominates until competition reaches a sufficient level such that the Schumpeterian effect prevails thereafter.[5] Their empirical evidence for manufacturing firms in the U.K. tends to support the hypothesis of an inverted-U pattern. Scherer (1967), Levin and Mowrey (1985) and Hashmi

non-financial sectors, Cetorelli and Strahan (2006) showed that potential entrants faced greater difficulty gaining access to credit in concentrated markets than in more competitive markets.

[2]The importance of financial services for economic growth has been long documented in the literature. Theoretical and empirical advances relate financial development to an improved allocation of capital, better risk sharing and possibly a higher savings rate (King and Levine, 1993; Pagano, 1993; Levine, 2004; Levine et al., 2000).

[3]For instance, Berger (2003) presents an overview of technical change in the US banking industry (e.g., automatic teller machines, electronic payments services, internet websites, information exchanges, computer-based credit risk scoring models, etc.) and infers that "competition may currently or in the near future force banks to adopt technology just to keep existing customers" (p. 149). He also observes that larger banks have been earlier adopters of new technologies than smaller banks. See also Mishkin and Strahan (1999), DeYoung and Hunter (2001), Berger and DeYoung (2006) and Jones and Critchfield (2005).

[4]See literature reviews by Kamien and Schwartz (1982) and Symeonidis (1996).

[5]Their model is also able to explain positive or negative effects of competition on innovation.

(2007) also find evidence of an inverted-U relationship between competition and innovation. Bos et al. (2009) extend the previous literature from manufacturing to financial services and find an inverted-U relationship between competition and innovation in U.S. banking.

An important limitation of the existing literature is the fact that it has ignored the effect of firm size on the competition-innovation relationship, relying instead on the premise that unit costs are independent from output levels (Aghion et al., 2005b). However, for many industries this assumption may not hold. An important example is the common finding of U-shaped average cost curves in the U.S. banking industry (Berger et al., 1999; Vives, 2001). In the presence of a Minimum Efficient Scale (MES), bank size may play an important role in the competition-innovation relationship, as some banks may become 'too big to innovate' : for these banks, the growth opportunities are limited compared to smaller firms below the MES or may even come at a penalty of higher average costs.[6] Consequently, the existence of scale diseconomies for these firms constitutes an additional reason for concern in a consolidating market, since these large firms may react to any further decreases in competition with strong decreases in their innovation effort, as was the case during the recent global financial crisis.[7]

The aim of this chapter is to examine this relationship between scale economies, competition and innovation in order to find out whether there is such a thing as being 'too big to innovate' in the U.S. banking sector in the period 1984-2004, a period of substantial consolidation. To our knowledge, this study constitutes the first theoretical and empirical investigation of the impact of scale (dis)economies on the competition-innovation nexus.

The chapter contributes to the literature in two distinct ways. The first contribution consists of an important extension of the theoretical model of Aghion et al. (2005b) to account for scale economies in studying the competition-innovation relationship. In the original model, all firms operate at identical, fixed average costs. As a result, innovation is independent

[6]Of course, for firms that operate below the MES, cost scale economies represent a bonus from innovation. These firms experience more growth opportunities compared to firms with scale diseconomies since firm growth results in lower average costs and increases in their profits. As a result, the existence of scale economies may constitute a key driver of consolidation.

[7]The large cuts on Information Communications Tecnology (ICT) spending by large banks during the recent financial crisis points in this direction. According to the Celent (2009) financial report, the growth in the information technology spending by the North American banks in 2009 is modest (1.7%), compared to past years' spending. The drop is even more dramatic for large European banks, who decreased ICT spending by 5.8% in 2009.

of firm size and scale economies. We augment the structure of the cost function to allow for U-shaped average costs, and modify the empirical model of Aghion et al. (2005b) to distinguish between firms with different scale (dis)economies.

The second contribution of this chapter lies in the application of a novel innovation measure in the context of this study. Instead of using traditional innovation measures based on outputs or inputs (e.g., patents and R&D spending, which are mostly relevant to manufacturing), we focus on banks' ability to minimize costs through process innovations. Following early work by Hayami and Ruttan (1970), Mundlak and Hellinghausen (1982) and Lau and Yotopoulos (1989), we estimate and envelope annual minimum cost frontiers to create a meta frontier, which represents the best (potential) available technology. The distance to the latter constitutes each bank's technology gap, which is reduced if the bank manages to innovate. Our proposed measure has three advantages. First, it enables us to examine the innovation behavior of firms in a sector where traditional measures such as patents, R&D expenditures, number of scientists and engineers are less applicable and suffer from several limitations (Kamien and Schwartz, 1982; Acs and Audretsch, 1987a; Geroski, 1990; Griliches, 1990).[8] Second, our measure closely aligns with the model of Aghion et al. (2005b) as each innovation leads to lower production costs (process innovation) and laggard firms can catch up with leaders by inventing and imitating. Thirdly, contrary to the past literature that focuses on one type of technology (e.g. adoption of ATMs), our proposed measure captures all types of invention that lead eventually to cost reductions.[9]

The theoretical implications of our model and application of our methodology are tested on a rich data set of U.S. banks. Our analysis is organized around the following questions: (i) How have scale economies and competition developed in the consolidating U.S. banking market? and (ii) How has the increase in bank size affected the relationship between competition and innovation in U.S. banking: have large U.S. banks become too big to innovate?

Our results are easy to summarize. We find evidence that many banks

[8]For example, not all innovations are patented and R&D expenditures may systematically understate the research activity of small firms. Frame and White (2004) argue that patents for financial products and services are not common, financial services firms rarely have R&D budgets and rarely employ scientists and engineers. See Frame and White (2004) for a discussion on the data restrictions in financial services concerning innovation measures.

[9]See, for instance, the work of Hannan and McDowell (1984).

in the U.S. experienced scale diseconomies during the consolidation period. In the same period, the upward trend in the average price cost margins implies that the degree of competition in the banking sector has declined. Furthermore, our results show that banks that operate above the MES have indeed become 'too big to innovate' : these banks' process innovation is more sensitive to changes in competition, confirming fears that further decreases in competition and more consolidation pose a serious threat to innovation in U.S. banking.

Our findings have important implications for competition policy. Although the current crisis has revived fears of banks having become too big to fail (Melvin and Taylor, 2009), the consolidation trend in U.S. banking continues. To some extent, charter values in banking may increase financial stability. But our results make clear that any further decreases in competition come at a high price: innovation is expected to fall sharply due to the strong reaction of large banks that have become 'too big to innovate' .

The remainder of the chapter is organized as follows. Section 3.2 describes how we extend the theoretical model of Aghion et al. (2005b) and its application to the banking sector. Section 3.3 discusses the data and methodology. Empirical results are presented in Section 3.4. Section 3.5 concludes.

3.2 Theoretical Framework

This section presents our model. The model is based on the theoretical contribution of Aghion et al. (2001) and Aghion et al. (2005b), who develop a growth model to investigate the relationship between competition and innovation. We follow Aghion et al. (2001) and Aghion et al. (2005b), unless it is stated otherwise. We start by explaining the set-up of the model, introducing a U-shaped average cost function. Next, we proceed with the derivation of equilibrium profits and the Schumpeterian and escape competition effect. Finally, we describe the effect of a U-shaped average cost function on the competition-innovation relationship.

3.2.1 Basic model with a U-shaped average cost function

In this subsection, we introduce the key elements of the model with a U-shaped average cost curve. We assume that a sector faces identical consumers with a constant inter-temporal discount rate, r, and a log-utility function that can be described by:

$$u(y_t) = \ln y_t, \tag{3.1}$$

where y_t denotes the consumption good of the sector.

The sector consists of a continuum of intermediate sectors that produce $y_t = \int_0^1 \ln x_{jt} d_j$, where x_{jt} is an aggregate of two intermediate goods produced by firm A and B (duopoly) in the intermediate sector j. The total production of each intermediate sector is $x_j = \sqrt[\alpha]{x_{Aj}^{\alpha} + x_{Bj}^{\alpha}}$, where α is the degree of substitutability between products.[10]

The innovation rates of a technological leader, laggard firms and neck-and-neck firms (that are at technological par with one another) are denoted n_1, n_{-1} and n_0, respectively. The laggard bank moves ahead with the hazard rate $n + h$ if it puts effort into R&D, where h is a help factor that represents R&D spillovers or the ability to copy the technology of a leader. The R&D cost function $\psi(n) = n^2/2$ is expressed in units of labor n. By assumption, technological advances occur through step-by-step innovations instead of leapfrogging and the maximum technological gap, m, in a sector is assumed to be one ($m = 1$) since laggard firms can adopt the leader's previous technology.[11] Therefore, $n_1 = 0$, as a firm that is already a leader has no incentive to innovate further. Firms (banks, in our case) only use labor as an input at the (exogenous) normalized wage rate $w = 1$ and it is assumed that the production function exhibits constant-returns.

A key assumption of the Aghion et al. (2005b) model is that unit costs are independent from the quantity produced. In particular, the unit cost function in their model is γ^{-k_i}, where γ represents the size of an innovation (and is assumed to be larger than one) and k is the technological level. Hence, innovations lower the unit cost due to a decrease in the required units of labor per unit of output. However, models based on the assumption that unit costs are independent from output levels are less appropriate in sectors where average costs and marginal costs are not constant and may thereby affect the competition-innovation relationship. In the model of Aghion et al. (2005b), for a given level of competition, innovation incentives depend on the incremental profits from innovation.[12] To see how non-constant average

[10] The logarithmic utility function in equation (3.1) implies that in equilibrium individuals spend the same amount on each aggregate of intermediate goods x_j.

[11] It is impossible for laggard firms to surpass a technological leader by means of an innovation without first drawing even with this leader. See Aghion et al. (1997) for several appealing features of a model of step-by-step innovation compared to the Schumpeterian leapfrogging models.

[12] The incremental profit from innovation is the difference between post- and pre-innovation profits.

costs affect this innovation incentive, consider the U-shaped average cost curves that exist in many sectors, such as the banking sector, where the MES is less than infinity. The existence of scale (dis)economies may influence the innovation incentives by affecting the incremental profits from innovation. For example, scale diseconomies lower current profits and can hamper the growth opportunities of firms and their future profits. However, in the model of Aghion et al. (2005b) firms cannot exhibit scale economies and consequently the current firm size and firms' growth potential cannot affect the innovation behavior of firms (through the relationship between scale economies and incremental profits from innovation).

We depart from Aghion et al. (2005b) and allow for U-shaped average costs. We propose the following total cost function:

$$TC_i = F + w\gamma^{-k_i} x_i^{\delta}, \tag{3.2}$$

where F are fixed costs, w is the wage rate (assumed to be equal to 1), δ is a non-negative integer and x_i is the output of firm i. The cost function in Aghion et al. (2005b) is a special case of equation (3.2), where $\delta = 1$ and $F = 0$. Since γ^{-k_i} represents the amount of labor needed to produce one output, $w\gamma^{-k_i}$ represents the labor expenses per unit of output. The average cost function is:

$$AVC_i = \frac{F}{x_i} + \gamma^{-k_i} x_i^{\delta-1}. \tag{3.3}$$

The domain of this average cost function is the open interval $x \in (0, \infty)$. Differentiating the function with respect to x_i and setting the derivative equal to zero, yields the output x_i^* that is associated with the Minimum Efficient Scale (MES) of production:

$$x_i^* = \sqrt[\delta]{\frac{F}{(\delta - 1)\,\gamma^{-k_i}}}. \tag{3.4}$$

Figure 3.1a illustrates the result, with firms producing below ($x_i < x_i^*$), at ($x_i = x_i^*$) or above ($x_i > x_i^*$) the MES. The optimal scale size of production depends on the size of an innovation and the technological level of a bank. Defining the unit labor requirement as $\Gamma = \gamma^{-k_i}$ and differentiating the MES with respect to the unit labor requirement shows that lower levels of unit labor requirements lead to a higher optimal scale:

$$\frac{\partial x_i^*}{\partial \Gamma} = \frac{1}{\delta} \left(\frac{F}{(\delta - 1)\,\gamma^{-k_i}} \right)^{\frac{1-\delta}{\delta}} \left(\frac{-F\,(\delta - 1)}{\left((\delta - 1)\,\gamma^{-k_i}\right)^2} \right) < 0. \tag{3.5}$$

Figure 3.1b illustrates this result, with the solid line depicting the average cost curve based on the current technology in period $t = 1$. After an innovation, the technology of the firm improves and the MES increases in period $t = 2$. Thus firms can operate at a higher optimal scale after improvements in their technology.

3.2.2 The escape competition and Schumpeterian effect

Aghion et al. (2001) show that the equilibrium profit of each firm depends only on its relative costs, z_i. Using the cost structure in equation (3.2), the relative unit costs (z_i) of a firm become:

$$z_i = \frac{MC_i}{MC_{-i}} = \frac{\delta\gamma^{-k_i} x_i^{\delta-1}}{\delta\gamma^{-k_{-i}} x_{-i}^{\delta-1}} = \frac{x_i^{\delta-1}}{x_{-i}^{\delta-1}}\gamma^{-m}, \tag{3.6}$$

where MC_i and MC_{-i} are the marginal costs of firms i and (the other firm) $-i$, respectively and m is the technological gap (lead or lag) of a firm.[13] If $\delta = 1$, we obtain the relative cost as in the model of Aghion et al. (2001), namely γ^{-m}.[14] The profit of a firm is a function of the relative marginal costs and the degree of substitutability between the products $\pi_i = f(z_i, \alpha)$.

We follow Aghion et al. (2005b) and model the degree of product market competition by the ability of two neck-and-neck firms to collude.[15] Firms are assumed to engage in Bertrand competition. The profits of laggards, neck-and-neck firms and technological leaders are π_{-1}, π_0 and π_1, respectively. Collusion is assumed to only be possible when firms have equal unit costs. The profits of firms with equal costs are defined as:

$$\pi_0 = \epsilon\pi_T, \quad \epsilon \in \left[\frac{\pi_{0_{nc}}}{\pi_T}, \frac{1}{2}\right], \tag{3.7}$$

where ϵ represents the share of the total (perfectly) collusive profits π_T, and $\pi_{0_{nc}}$ represents the profits when both neck-and-neck firms do not collude. The share of the maximum total collusive profits ϵ is exogenously determined in the model and defined on a closed interval, where π_{0nc}/π_T is the minimum

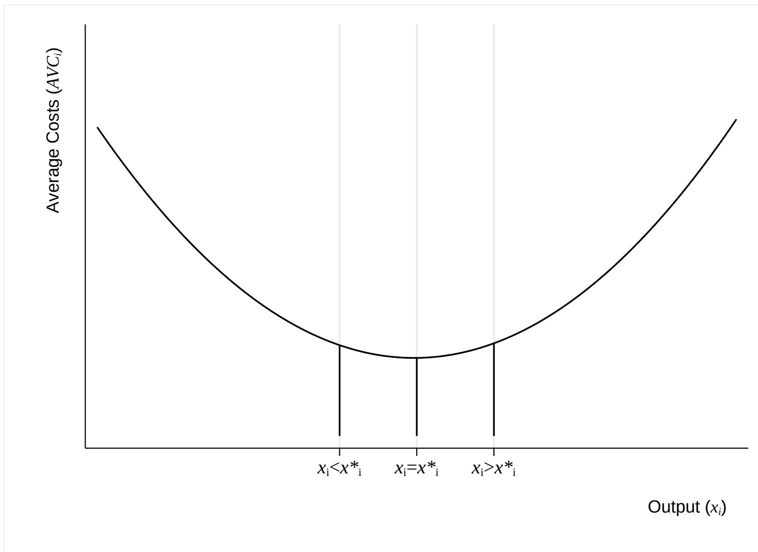
[13] Firms are assumed to be homogeneous with respect to the δ parameter.

[14] Profits also depend on the degree of substitutability between the two products of the two intermediate sectors. Since the relative marginal cost function is not defined if x_{-i} is zero, we assume that the degree of substitutability is not perfect and hence firms always produce.

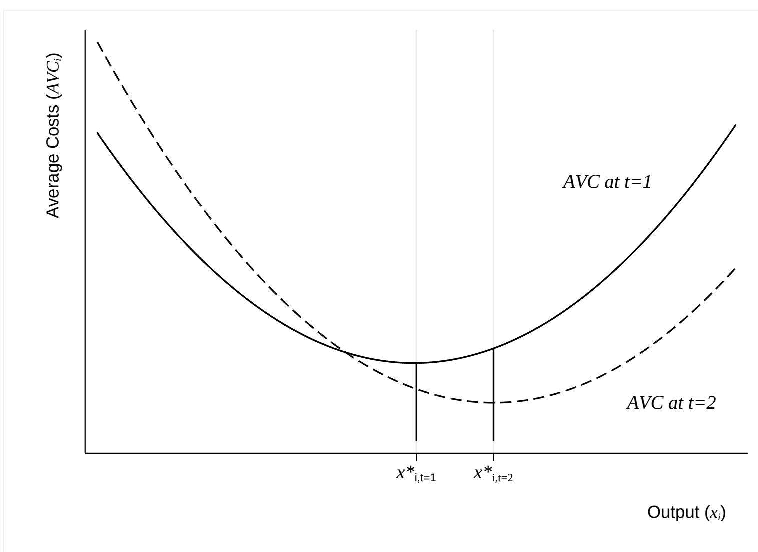
[15] Aghion et al. (2001) use the substitutability between products parameter α as a competition measure.

Figure 3.1: U-shaped average costs, optimal size and innovation

(a) Operating above, at and below the minimum efficient scale (MES)



(b) Average costs for different values of the unit labor requirement



share value that indicates no collusion and the maximum share value 1/2 indicates an equal division of the total (perfectly) collusive profits.[16] From equation (3.7), we learn that the profits of neck-and-neck firms π_0 can be used as a measure of competition. Ceteris paribus, more collusion leads to higher profits and less collusion to lower profits.

As in Aghion et al. (2005a), the following equilibrium research intensities, n_0 and n_{-1}, are obtained from the Bellman equations:

$$n_0 = -h + \sqrt{h^2 + 2(\pi_1 - \pi_0)}, \tag{3.8}$$

$$n_{-1} = -(h + n_0) + \sqrt{(h + n_0)^2 + 2(\pi_0 - \pi_{-1}) + n_0^2}. \tag{3.9}$$

The research intensities depend only on the current technological state. Furthermore, the discount rate is assumed to be zero ($r = 0$). These research intensities can be used to examine how the escape competition and Schumpeterian effect are affected if firms experience scale economies or scale diseconomies. The escape competition and Schumpeterian effect are obtained by differentiating the research intensities with respect to the measure of competition π_0. Equation (3.10) shows the escape competition effect:

$$\frac{\partial n_0}{\partial \pi_0} = -\frac{1}{\sqrt{h^2 + 2(\pi_1 - \pi_0)}} < 0. \tag{3.10}$$

Equation (3.10) illustrates that competition, reflected by a lower π_0, has a *positive* effect on the innovation incentives of neck-and-neck firms. More competition increases incremental profits from innovating by lowering the current profits. Equation (3.11) represents the Schumpeterian effect:

$$\frac{\partial n_{-1}}{\partial \pi_0} = -\frac{\partial n_0}{\partial \pi_0} + \frac{((h + 2n_0)\frac{\partial n_0}{\partial \pi_0} + 1)}{\sqrt{(h + n_0)^2 + 2(\pi_0 - \pi_{-1}) + n_0^2}} > 0. \tag{3.11}$$

It shows that competition has a *negative* effect on the innovation incentives of laggard firms. Increases in competition reduce the rents that these laggard firms can earn after they innovate and hence lead to a decrease in incremental profits from innovating. This decrease in their incremental profit lowers their innovation incentives.

[16] Tirole (1988) argues that the effect of decreasing returns to scale on collusion is ambiguous. Undercutting may become less profitable with decreasing returns to scale but also weakens the strength of retaliation.

Aghion et al. (2005b) show that the inverted-U relationship between competition and innovation depends on the escape competition effect and the Schumpeterian effect. Below the optimal level of competition that enhances innovation, the escape competition effect dominates and beyond the optimal level of competition the Schumpeterian effect dominates. In the next section, we extend the model of Aghion et al. (2005b) to examine the effect of scale (dis)economies on the escape competition effect, the Schumpeterian effect and the competition-innovation relationship.

3.2.3 The effect of a U-shaped average cost function on the competition-innovation relationship

Producing below or beyond the MES may lead to differences between the incremental profits from innovating. Post innovation profits are affected by scale economies in two ways. First, after innovating firms receive a bonus (pay a penalty) since, operating below (above) the MES, leads to more (less) growth potential. Firms with diseconomies of scale have less growth potential since firm growth may lead to more than proportional increases in their costs. Second, *differences* in scale economies can have an indirect effect, by changing the likelihood of collusion.[17] Scale diseconomies make undercutting a less profitable strategy in the present due to increases in average costs. However, scale diseconomies may also foster competitive behavior by reducing the capacity to retaliate against competitive pricing. As a result, it is as yet unclear how a U-shaped average cost curve affects the ability to collude (Tirole, 1988).

However, we can investigate the direct effect of scale economies. To start with, the relationship between scale economies and the escape competition can be examined by differentiating equation (3.10) with respect to the incremental profits $\chi = \pi_1 - \pi_0$:

$$\frac{\partial^2 n_0}{\partial \pi_0 \partial \chi} = \frac{1}{(h^2 + 2\chi)^{3/2}} > 0. \tag{3.12}$$

Equation (3.12) is used to examine the magnitude of the escape competition effect if incremental profits (χ) of neck-and-neck firms change. Higher incremental profits result in a lower escape competition effect. Neck-and-neck firms with scale economies have higher post-innovation rents π_1 if their average costs decrease due to firm growth. Large neck-and-neck firms with

[17] Although we discuss the possible links between scale economies and collusion in this chapter, competition is not endogenously determined in this chapter and Aghion et al. (2005b).

scale diseconomies have lower incremental profits χ if they expect lower post-innovation rents π_1 and/or higher current profits π_0 due to collusive behavior. Consequently, these large firms experience a larger escape competition effect.[18]

Differentiating equation (3.11) with respect to the gap between the profit of a neck-and-neck and laggard bank ($\lambda = \pi_0 - \pi_{-1}$), shows how the Schumpeterian effect is affected by this gap in profits:

$$\frac{\partial^2 n_{-1}}{\partial \pi_0 \partial \lambda} = -\frac{(h + 2n_0)\frac{\partial n_0}{\partial \pi_0} + 1}{((h + n_0)^2 + 2\lambda + n_0^2)^{3/2}} > 0. \tag{3.13}$$

The Schumpeterian effect is larger if the incremental profits (λ) from innovation of a laggard firm are higher. Hence, laggard firms with scale economies are more likely to innovate, for a given level of competition. Scale economies affect the incremental profit $\lambda = \pi_0 - \pi_{-1}$ since only laggards are subject to the Schumpeterian effect. Post-innovation rents of laggards π_0 with scale diseconomies may be higher if these large firms are more likely to collude.[19] The initial profits of laggards π_{-1} with scale diseconomies may be lower compared to firms on the MES, since scale diseconomies may suppress the profits that can be earned. Therefore, laggard firms with scale diseconomies that are more likely to collude after they innovate and experience suppressed current profits have higher incremental profits (λ) from innovation and a larger Schumpeterian effect.[20]

In sum, the magnitude of the escape competition and Schumpeterian effect depends on the current level of the incremental profits, which in turn are affected by the scale (dis)economies of a bank. The influence of scale economies on incremental profits is ambiguous and hence an empirical issue. The slope of the inverted-U relationship also depends on the magnitudes of the escape competition and Schumpeterian effect. If banks indeed become 'too big to innovate', larger escape competition and Schumpeterian effects may lead to a steeper inverted-U relationship between competition and innovation.

[18]The escape competition effect is only smaller if firms with scale diseconomies are more likely to behave competitively (low π_0) and if this competitive behavior results in high incremental profits χ from innovation.

[19]However, the reverse is also possible, since the effect of scale diseconomies on collusion is ambiguous.

[20]The Schumpeterian competition effect is only smaller if laggards with scale diseconomies are more likely to behave competitively and experience increases in their average costs (low π_0). These laggards have low incremental profits and are relatively less affected by an increase in competition.

3.3 Data and Methodology

This section provides a short description of the data and presents innovation, scale economies and competition measures, as well as our set of control variables and our estimation strategy.

3.3.1 Data

Our sample consists of a large number of individual banks over the period 1984-2004 in the United States. On average, around 10,500 banks per year are included in the dataset.[21] Information is gathered from the Call reports for Income and Condition provided by the Federal Reserve System. The Call Reports are quarterly income statement and balance sheet data that all federally insured banks are required to submit to the Federal Deposit Insurance Corporation (FDIC). The data cover all banks regulated by the Federal Reserve System, the Federal Deposit and Insurance Corporation and the Comptroller of the Currency. For the purpose of this study, we include only independent banks, in order to avoid measurement problems, in particular when measuring competition.[22] The Call Reports include complete balance sheet and income statement data. All data are expressed in 1984 U.S. dollars.

3.3.2 Measuring Innovation

In Aghion et al. (2001) and Aghion et al. (2005b), innovations result in changes in the unit labor requirement, as shown in Figure 3.1b in the previous section. In the empirical literature, relating shifts in a cost function, i.e., technical change, to (process) innovation is not new (Subramanian and Nilakanta, 1996; Ruttan, 1997; Agrell et al., 2002; Bleaney and Wakelin, 2002). Our approach to measuring innovation falls into this tradition, and closely aligns with the theoretical concept of the technology gap as stipulated in Aghion et al. (2001) and Aghion et al. (2005b).

The measurement of technical change has gone through a number of developments. For a long time, the econometric approach, led by Tinbergen (1942), consisted of including a time trend when estimating a cost (or production) function. Likewise, using index number theory, Solow (1957) identified neutral technical change, with constant marginal rates of substi-

[21] We assume that attrition of the panel data occurs exogenously.

[22] We selected independent banks using Call Report items RSSD9397, RSSD9001 and RSSD9365.

tutions. Put differently, in line with Tinbergen (1942) and Solow (1957), a change in the cost curve in Figure 3.1b would be a parallel shift, leaving the (other) parameters in the cost function unaffected.

For the index numbers approach, Diewert (1976) and others that followed, added flexibility to the measurement of technical change, relaxing the assumption that the latter was constant. For the econometric approach, Baltagi and Griffin (1988), by introducing a general index of technical change, relax the assumption that technical change is constant in time. In addition, building on advances by Gollop and Jorgenson (1980) and others, they allow for biased technical change, as marginal rates of substitution are allowed to vary over time. The result is depicted in Figure 3.1b, where a shift of the cost curve from $t = 1$ to $t = 2$ is not (necessarily) parallel, and can, for example, reflect a skill bias.[23]

The advances described so far, consist of adding flexibility to the aggregate cost function, which is thereby allowed to evolve over time. Since the cost function describes the process whereby firms are assumed to minimize costs while producing output, the set of estimated parameters reflects the state of technology. What separates technical change from other means of minimizing costs (such as adjusting the input mix or reducing waste), is the fact that whereas the latter measures use the currently implemented technology, technical change consists of the invention or adoption of a new technology. In an attempt to reconcile this view of technical change with the notion of estimating cost functions, early work by Hayami and Ruttan (1970), Mundlak and Hellinghausen (1982) and Lau and Yotopoulos (1989) has focused on the notion of a *meta frontier*. The latter encompasses the set of available technologies, across firms and/or across time. In this view, technical change consists of the adoption of a new technology, and is measured against the benchmark meta frontier, which combines all available technologies.

Figure 3.2a is used to explain the concept of the meta frontier, with a simple example based on cost minimization with two inputs (x_1, x_2) and a single output (y) and for two firms I and II. In this example, there are two annual frontiers, for $t = 1$ and time $t = 2$. Each frontier represents the minimum cost curve for a certain level of output, based on the available technology in period t. The cost efficiency of firm I, located at point E at time $t = 1$ is OD/OE. If firm I is at point C at $t = 2$, its efficiency is OB/OC. Figure 3.2a also shows that a second firm, II, is located at point

[23]The relationship between biased technical change and innovation is explained in Acemoglu (2002).

H and faces a cost efficiency of OG/OH at time $t = 1$. The dashed line that envelops the annual frontier in Figure 3.2a represents the minimum cost frontier over the whole period, or meta frontier. Innovation results in a lower gap between the annual minimum cost frontier and the meta frontier. Innovation is then reflected by changes in the technology gap, which measures the difference between currently available technology and optimal, available technology over the whole period with values between zero and one (i.e., the firm is on the meta frontier).[24] At $t = 1$ the firm faces a technology gap of OA/OD, which narrows to OA/OB at $t = 2$ as the firm improves its technology set. While firm I faces a technology gap of OA/OD in period $t = 1$, the technology gap of firm II is smaller (OF/OG).

We follow recent work by Battese et al. (2004), O'Donnell et al. (2008) and Bos and Schmiedel (2007), and obtain technology gaps by, first, employing Stochastic Frontier Analysis (SFA) to estimate the minimum cost frontier available in each year, and second, enveloping the resulting annual cost frontiers to obtain a meta frontier by following early works of Hayami and Ruttan (1970), Mundlak and Hellinghausen (1982) and Lau and Yotopoulos (1989).[25]

The first step consists of estimating an annual translog cost frontiers using stochastic frontier analysis.[26]

$$TC_{it} = f^*(w_{it}, y_{it}, z_{it})_{it} e^{v_{it}+u_{it}}, \tag{3.14}$$

where w represents the vector of input prices, y the output vector, z a vector of control variables, v the random noise component which is assumed to be i.i.d. $N(0, \sigma_{vit})$, and u the inefficiency term which is assumed to be i.i.d. $N|(\mu_{it}, \sigma_{uit})|$. We use stochastic frontier analysis, since this allows for the presence of inefficiency, which is completely ignored by the conventional measures of productivity (e.g., TFP) that conflate technical change with efficiency change (Bos et al., 2009).

We assume that banks minimize their total costs and operate in perfectly competitive input markets. The activities of the banks are specified according to the so-called intermediation approach (Freixas and Rochet, 1997). Therefore, the output vector y consists of loans, investments, and

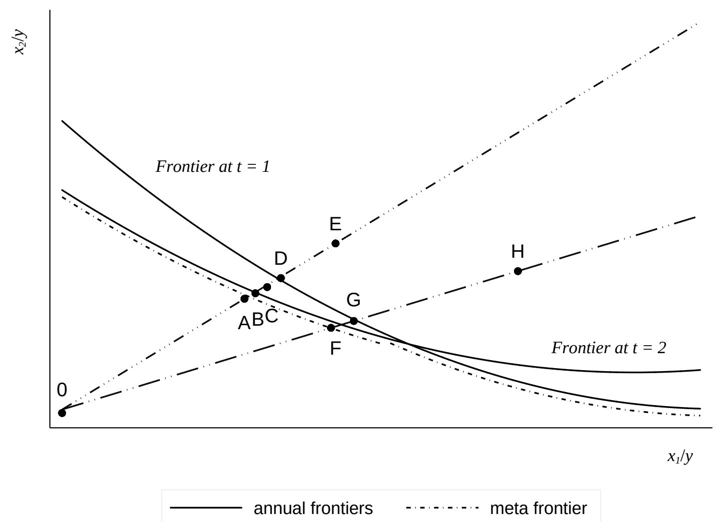
[24] The notion of technology gaps has been first used by Krugman (1979) and proxied in the literature by total factor productivity (TFP) differentials (Griffith et al., 2004).

[25] Kumbhakar and Lovell (2000) provide an elaborate discussion of the development and application of SFA to efficiency measurement.

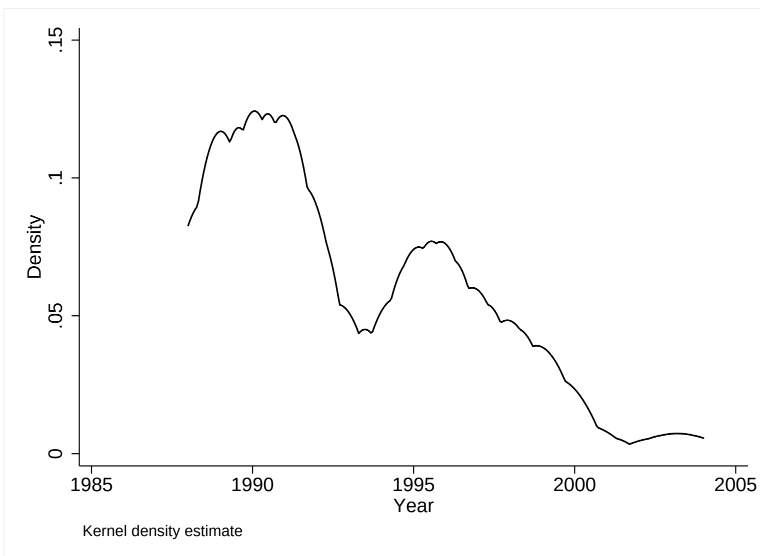
[26] Homogeneity of degree one in input prices and symmetry are imposed.

Figure 3.2: Technolog gaps in U.S. banking

(a) Meta frontier and technology gap



(b) Distribution of banks with a technology gap of 1



off-balance sheet items.[27] Input prices, w, include the price of fixed assets, the price of labor, and the price of borrowed funds.[28] The equity ratio is included as a control variable z to account for different risk profiles of banks (Hughes and Mester, 1993). The composed error in equation (3.14) is $\varepsilon_{it} = \nu_{it} + u_{it}$. Firm-specific inefficiency estimates $u_{i,t}$ are obtained by using the expected value of u_{it} conditional on the total error ε_{it} (i.e., $E(u_{it}|\varepsilon_{it})$). Cost efficiency score estimates are obtained as follows:

$$C\hat{E}_{it} = [\exp(-\hat{u}_{it})], \tag{3.15}$$

where CE equals 1 for banks that operate on the annual frontier (no inefficiency). Banks that are subject to inefficiencies are operating above the annual cost frontier and have cost efficiency scores less than 1.

Our second step consists of estimating the meta frontier as the envelope around these annual cost frontiers. We use the parameter estimates for the annual cost frontiers and estimate the distance between the annual frontier (f^*) and meta frontier (f_{meta}):

$$\begin{aligned} Min.\text{Distance}_{it} = \sum_{t=1}^{T}\sum_{i=1}^{N} |lnf^*(w_{it}, y_{it}, z_{it})_{it} - lnf_{meta}(w_{it}, y_{it}, z_{it})_{it}| \\ s.t.\ lnf_{meta}(\cdot)_{it} \leq lnf^*(\cdot)_{it}. \end{aligned} \tag{3.16}$$

In the constrained minimization problem above, the absolute distance between the annual cost frontier and the meta frontier is minimized subject to the constraint that the total costs from the annual frontier is equal to or larger than the total costs from the meta frontier. As a result, the technology gap is defined as:

$$GAP_{it} = \frac{f_{meta}(w_{it}, y_{it}, z_{it})_{it}}{f^*(w_{it}, y_{it}, z_{it})_{it}}. \tag{3.17}$$

Innovations by firms may lead to improvements in their technology set and therefore a smaller gap between the current technology set and the (potentially available) best technology set, namely the meta frontier. The

[27]The output quantities are year-end stocks.

[28]The price of fixed assets is calculated by taking the depreciation over fixed assets, the price of labor is obtained by dividing the personnel expenses by the number of fte-employees and the price of borrowed funds equals the interest expense divided by the total borrowed funds.

result is an increase in the GAP_{it}, which is bounded between 0 and 1, where the latter is reached when firms operate on the meta frontier.

The mean of the technology gap in our data is 0.989 (see Table 3.1 on page 58). Figure 3.2b shows the distribution of banks over the period 1984-2004 with a technology gap equal to 1. This means that these banks are using the best of the available technologies over the period and operate on the meta frontier. The average bank size was much smaller in the 1980s and early 1990s. This may be one explanation why, in this period, banks had an easier time reaching the meta frontier, given the amount of inputs and outputs involved in the production process. As banks increased in size, and presumably in complexity, bridging the technology gap can become more challenging. Another explanation for the same trend, however, may be a possible decline in competition during the sample period. We therefore now turn to our approach to measuring competition.

3.3.3 Measuring Competition

We measure competition in the banking sector by using one minus the price cost margin (viz., the Lerner index, or markup) as in Aghion et al. (2005b):

$$C_{it} = 1 - \left(\frac{\Pi_{it} + F_{it}}{R_{it}}\right), \tag{3.18}$$

where Π_{it} is the profit of a bank, F_{it} represents the fixed costs and R_{it} are the total sales of each bank. The price cost margin of a bank is obtained by dividing the net income after taxes and extraordinary items plus expenses of premises and fixed assets by total non-interest income plus total interest income. A firm-specific measure of competition is preferred instead of using a measure based on a certain geographical market as some banks compete mainly at the local level while other banks compete mainly at the national or international level. Hence, we assume that all changes in competition are reflected in the price cost margins of banks despite of the scope of the geographical market in which banks are competing.[29] After the exclusion of outliers, the price cost margin and the competition measure range between -1 and 1, and 0 and 2, respectively.[30]

[29] An important assumption in this thesis is that less competition is positively related with higher price cost margins, while more competition leads to lower price cost margins.

[30] Scatterplots of the technology gap and the price cost margin were used to check for outliers. In total, 1324 observations (less than 1% of the original sample) were excluded. A range of the price cost margin between -1 and 1 was considered to be reasonable. Some authors choose to remove negative price cost margins, but this approach creates a bias

3.3.4 Measuring Scale Economies

We employ two measures for scale economies. The first measure is a firm size measure based on total assets.[31] There should be a direct correlation between scale economies and firm size as firms that are below the MES are usually the (relatively) smaller firms in an industry. This measure, however, has two main drawbacks. First, the firm size variable may capture other effects than scale economies in production. Second, while a certain firm size may be characterizing the MES in a given year, the same firm size may indicate scale economies in other years (see Figure 1).

An alternative measure is based on the elasticity of total costs with respect to the outputs. By using SFA, we estimate cost functions and calculate the elasticity as follows:[32]

$$Scale_{it} = \sum_{k=1}^{3} \frac{\partial \ln TC_{it}}{\partial \ln y_{kit}}, \tag{3.19}$$

where k indicates the three different outputs used in this chapter, namely loans, investments and off-balance sheet items.

The SFA scale economies are calculated for a given output bundle keeping the output mix constant.[33] If the scale economies variable is equal to 1, the production is subject to constant returns to scale. If the value is lower than 1, it indicates scale economies. Values larger than 1 indicate scale diseconomies. A drawback of this measure is that it is a generated regressor that is constructed from a similar underlying procedure that is used to obtain the technology gap.

3.3.5 Control variables

We include two control variables in our analysis. The equity ratio is included as a control variable to account for the relationship between the risk of a

in the results as only firms with positive price cost margins are considered. The dataset contains 14,073 observations with negative price cost margins.

[31] The hypothesis that large firms are more than proportionally innovative compared to smaller firms is associated with the work of Schumpeter. There are many channels through which larger firms may have innovation advantages over smaller firms. For example, larger firms may have an advantage because they have more researchers which may lead to more productive interaction in this large research group. It is also possible that larger firms are more able to diversify risky innovation projects.

[32] For example, see Hunter and Timme (1991), Bernstein (1996) and Altunbas et al. (2001) for applications of this scale economies measure in the banking sector.

[33] As opposed to scale biased technical change, where technical change may result in a change of the cost-minimizing size of the firm (Hunter and Timme, 1991).

firm and innovation. Aghion et al. (2005b) argue that more debt pressure has a positive effect on the innovation incentives of firms. This positive effect is interpreted as an attempt of firms to escape from their existing debt pressure through innovations. Furthermore, the average wage per fte-employee (salary expenses divided by the number of fte-employees) is also included in our analysis as a proxy for human capital, as high quality workers may foster innovations (Funke and Strulik, 2000).[34]

Table 3.1: Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.
Technology gap	0.989	0.029	1.07e-08	1 .000
Price cost margin	0.179	0.090	-0.993	0.964
Total assets per millions of USD	458.740	7416.978	1.067	967,365
Scale economies (from SFA)	1.099	0.071	0.726	1.737
Risk (Equity/Total Assets)	0.096	0.034	7.36e-05	0.998
Salary expenses per fte in thousands of USD	35.143	12.756	0.048	537.160

151,476 observations. The descriptive statistics are based on the sample of the preferred specifications in Table 3.2.

Table 3.1 provides the descriptive statistics of the variables employed in our analysis.[35] We observe a large spread in the price cost margin. Also, both scale economies (values smaller than unity) and scale diseconomies exist in our sample.

3.3.6 Empirical Specification

Our empirical specification is based on Aghion et al. (2005b), but extended to account for the impact of scale (dis)economies:

$$GAP_{it} = \beta_1 C_{it} + \beta_2 C_{it}^2 + \beta_3 S_{it} + \beta_4 S_{it} C_{it} + \beta_5 S_{it} C_{it}^2 + \gamma' \mathbf{Z}_{it} + a_i + \varepsilon_{it}, \quad (3.20)$$

where GAP is the technology gap, C competition, S total assets or the SFA scale economies variable, γ' 1xn parameter vector, and $\mathbf{Z}$ an nx1 vector of control variables. A squared term of the competition variable, C^2, is

[34]We assume that the wages of high quality workers are higher.

[35]The total number of observations is approximately 220,000 with on average 10,500 banks each year. Eventually, we have 151,476 observations in the preferred specification due to missing values of several variables, the exclusion of outliers, applying first-differences, and using lags of the endogenous regressors as instruments.

included to account for the possible inverted-U relationship between competition and innovation according to Aghion et al. (2005). The interaction terms with the scale economies measure, SC and SC^2, are included to allow the inverted-U relationship to be different for each firm due to differences in scale economies.

The conditions in equation (3.12) and (3.13) show that the magnitude of the escape competition effect and Schumpeterian effect depends on the incremental profits that a firm can earn by innovating. Whether a firm experiences scale economies or scale diseconomies may affect this incremental profit and therefore the magnitude of the escape competition effect, Schumpeterian effect and the steepness of the inverted-U relationship. Taking first-differences of equation (3.20) to eliminate the unobserved heterogeneity a_i gives:

$$\begin{aligned}\Delta GAP_{it} = \beta_1 \Delta C_{it} + \beta_2 \Delta C_{it}^2 + \beta_3 \Delta S_{it} + \beta_4 \Delta (S_{it} C_{it}) + \beta_5 \Delta (S_{it} C_{it}^2) \\ + \gamma' \Delta \mathbf{Z}_{it} + \Delta \varepsilon_{it}.\end{aligned} \tag{3.21}$$

We also estimate equation (3.21) replacing the continuous scale economies measure with a dummy variable that indicates above average scale economies (1) or below average scale economies (0).

While competition and scale economies affect innovation, innovation may also affect competition and scale economies. For example, firms may become more dominant in a market after surpassing other competitors due to successful innovations.[36] Innovations may also increase the MES. Therefore, lags of these endogenous variables are used as instruments. The lag structure of the instruments depends on the order of autocorrelation in the residuals.[37] If the residual in equation (3.20) is not autocorrelated, lags of period *t-2* can be used as instruments for the endogenous regressors in equation (3.21). If there is first-order autocorrelation, lags from period *t-3* and deeper can be used. Furthermore, we use the two-step generalized method of moments estimator (GMM). The GMM has some efficiency gains compared to the traditional instrumental variable (IV) or two-stage least

[36] However, Geroski and Pomroy (1990) argue that more innovations may lead to more competition. Chapter 4 examines the effect of innovation output and innovation input on competition.

[37] The Arellano-Bond autocorrelation test is based on the examination of residuals in first differences. Testing for first-order serial correlation in levels is based on testing for second-order serial correlation in first differences.

square (2SLS) estimators. For example, the two-step GMM estimator utilizes an optimal weighting matrix that minimizes the asymptotic variance of the estimator. Also, GMM is more efficient than the 2SLS estimator in the presence of heteroskedasticity.

3.4 Results

In this section, we present our results. First, we investigate the development of the scale economies, average bank size and the average price cost margin over time. Next, we explore whether U.S. banks have become too big to innovate.

3.4.1 How have scale economies and competition developed in the consolidating U.S. banking market?

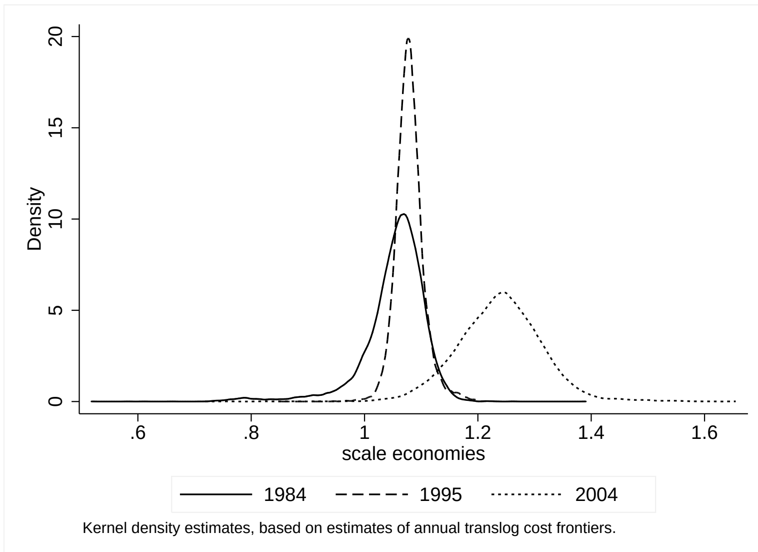
The model of Aghion et al. (2001) and Aghion et al. (2005b) analyzes how firms' innovative behavior changes as their competitive position changes. The resulting market dynamic is therefore grounded in firm-level differences. If there is no room to explore scale economies, banks are less able to reap the bonus that awaits them when they innovate. Likewise, if there are no banks operating beyond their MES, there is no such thing as being 'too big to innovate'.

Figure 3.3a shows how the distribution of scale economies has evolved over our sample period. The figure shows two concurrent developments. First, the number of banks with scale diseconomies has grown over time. Second, as the market consolidated, the distribution of scale economies narrowed mid sample (i.e., in 1995), but widened again towards the end of the sample period. Summing up, Figure 3.3a reflects a consolidating market, with plenty of differences in the potential penalty (or bonus) for innovators.

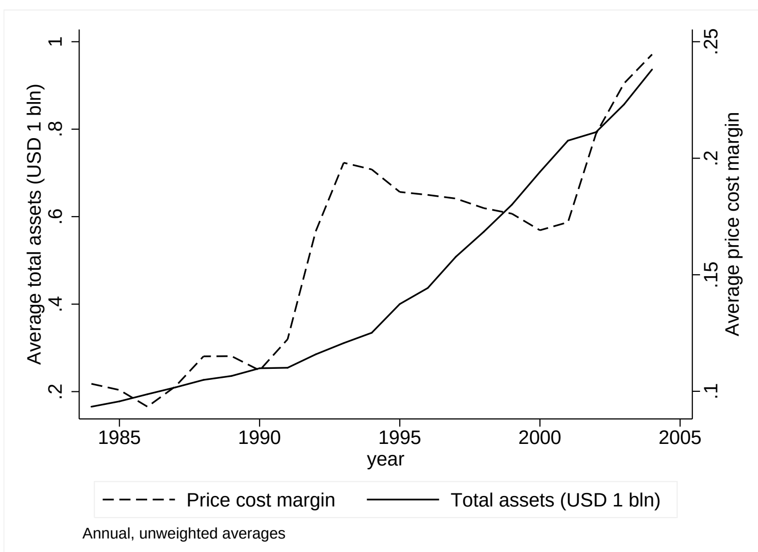
The story of a consolidating market is repeated in Figure 3.3b, which shows the development of average bank size (total assets in $bn) and the average price cost margin over the sample period. Interestingly, the peak in the price cost margin more or less coincides with the mid sample period, when the market was at its most homogenous in terms of scale economies. As Figure 3.3b shows, competition has decreased, on average, over the sample period. This finding is consistent with the results of Stiroh and Strahan (2003), who argue that the banking deregulation ignited a reallocation of banking assets from low profit to high profit banks. Meanwhile, the average

Figure 3.3: Scale economies, total assets and the price cost margin

(a) Distribution of scale economies in 1984, 1995 and 2004



(b) Average total assets and the price cost margin



bank size increased from total assets around 165 million dollars to around 936 million dollars in 2004.

The developments in scale economies and competition, together with the development of the technology gap described in Figure 3.2b, tell the story of a market with a lot of dynamics, and plenty of potential for banks to become 'too big to innovate'. The latter issue is investigated further in the next subsection.

3.4.2 Have large US banks become too big to innovate?

In this section, we investigate the influence of scale economies on the relationship between competition and innovation. Before we examine whether large US banks have indeed become 'too big to innovate', we need to establish the overall nature of the relationship between competition and innovation.

We therefore start with the first set of estimation results in Table 3.2, where we estimate the basic model specification, including control variables and total assets, but without allowing for a possible effect of the latter on the competition-innovation relationship. The results are shown in columns (i) and (ii), where we report both OLS (in column (*i*)) and two-step GMM (in column (*ii*)) estimates.[38] As is clear from the significant positive (negative) relationship for the competition measure (squared competition measure), there is indeed an inverted-U competition-innovation relationship: as competition increases, innovation initially rises, until it reaches a peak, after which the Schumpeterian competition effect dominates.[39] Interestingly, the optimal price cost margin that enhances innovation is 5.8%. Looking back at Figure 3.3b, we observe that on average this price cost margin is first reached around 1992, and again around the year 2000. Of course, comparing with Figure 3.3a, we know that a major difference between these two (sub-)periods lies with the development of scale economies.

[38]For the GMM estimation, we use instruments from periods $t-3$ and $t-4$, since there is evidence of first-order autocorrelation in the residuals in levels. The instruments are relevant and valid according to first-stage F-tests and the Hansen test, respectively. The instrument relevance test is based on the joint significance of the instruments, with a critical value for the F-statistic of 10 applied as an often used rule of thumb. The F-statistics for the regressions with the competition variable and its squared term as a dependent variable are 413.84 and 251.56, respectively.

[39]Both competition variables are individually and jointly significant at the 1% level.

Table 3.2: Estimation results: inverted-U

Specification	(*i*) *OLS*	(*ii*) *2-step GMM*	(iii) *2-step GMM*	(iv) *2-step GMM*	(v) *2-step GMM*
ΔC_{it}	0.025***	1.474***	1.468***	0.863**	0.303
	(0.009)	(0.106)	(0.105)	(0.348)	(0.847)
ΔC_{it}^2	-0.022***	-0.783***	-0.778***	-0.458**	-0.068
	(0.005)	(0.062)	(0.062)	(0.183)	(0.472)
ΔS_{it}	-0.0001**	-0.00009	-0.006	-0.281	-1.663***
	(0.00006)	(0.00006)	(0.005)	(0.256)	(0.522)
$\Delta S_{it} * C_{it}$			0.018	0.556	4.331***
			(0.012)	(0.500)	(1.240)
$\Delta S_{it} * C_{it}^2$			-0.013	-0.360	-2.720***
			(0.009)	(0.275)	(0.726)
ΔR_{it}	-0.162***	-0.124***	-0.123***	-0.107***	-0.749***
	(0.015)	(0.037)	(0.037)	(0.032)	(0.131)
ΔW_{it}	-0.0001**	-0.00006	-0.00006	0.00009	-0.00006
	(0.00004)	(0.00006)	(0.00006)	(0.00006)	(0.00009)
Observations	198,785	151,476	151,476	151,476	150,537
AR(1)		0.000	0.000	0.000	0.000
AR(2)		0.000	0.000	0.000	0.000
AR(3)		0.795	0.770	0.4421	0.8615
Hansen J statistic		0.797	3.609	14.417	8.671
		(0.671)	(0.607)	(0.0132)	(0.013)

The dependent variable is innovation, proxied by the technology gap, GAP. C is competition, S is scale economies proxied by total assets (or estimated scale economies, in specification (v)), R is the equity ratio, and W is the average wage per fte-employee. Standard errors (between parentheses) are robust against heteroskedasticity and serial correlation. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01. The p-values are reported for the Arellano-Bond serial correlation test. The chi-squared statistic and p-value (between brackets) are reported for the Hansen test.

These results are in line with Bos et al. (2009), who also report an inverted-U relationship for U.S. banking. In line with Aghion et al. (2005b), we find that banks with a higher equity ratio feel less debt pressure and are as a result less innovative. In our (preferred) GMM specification, a higher average wage has no effect, reflecting the fact that human capital has no considerable effect on banks' (process) innovativeness. And although firm size (S) is statistically significant at the 5% level in column (i), there is no evidence of a significant relationship between firm size and the technology gap once we treat the competition variables as endogenous regressors (column (ii)).[40]

Next, we proceed by investigating whether large U.S. banks have become to 'too big to innovate'. To this purpose, we allow interaction between scale economies and competition, so we can measure the effect of scale economies on the competition-innovation relationship. To accurately study this effect, we include scale economies in three different ways, presented in columns (iii), (iv) and (v) of Table 3.2. Each approach has its own advantages. In the first approach, presented in column (iii), scale is accounted for by banks' total assets. The advantages of this scale measure are that it is continuous, and (presumably) precisely measurable.[41] Therefore, we shall use this estimation in comparison with the earlier described results, in order to assess the effect of scale on the optimal price cost margin. In the second approach, presented in column (iv), we replace the scale measure with a dummy, which takes on the value 1 (0) when bank size is above (below) the industry average.[42] The advantage of this scale measure is that it allows us to easily compare the effect of operating below and above average total assets on the shape of the inverted-U relationship. Our third and final approach consists of measuring scale with scale economies estimates, as described in Table 3.1. Based on these estimates, we define a dummy that takes on a value of 1 (0) if a bank has scale economies above (below) the sample average.[43] The advantage of

[40] There is also no significant effect of firm size on the technology gap if we treat the firm size variable as endogenous with lags from period t-3 and t-4 as instruments. In the remaining part of the chapter we do not examine the marginal effect of firm size on the technology gap and focus only on the (conditional) marginal effect of competition on the technology gap.

[41] A possible disadvantage is that other things besides scale economies may affect the competition-innovation relationship.

[42] A possible disadvantage of this measure is the (lack of) accuracy with which we measure MES. We have also experimented with other cut-off points, including those based on our estimated scale economies (from Table 3.1). The results are robust.

[43] We also estimated column (v) using scale economies of unity as a cut-off point. Results are qualitatively similar.

this measure is that it allows us to more accurately differentiate between banks operating above and below average scale economies.[44] Comparing coefficients in columns (*iii*), (*iv*) and (*v*) reveals that results in all cases are qualitatively similar. The inverted-U relationship and the joint effect of the interaction terms are significant in all three cases.[45]

We start by evaluating the effect of scale on the optimal price cost margin. To do so, we compare the results in column (*ii*) with those in column (*iii*).[46] Our findings suggest that the inverted-U relationship between competition and innovation is bank-specific, and depends on the size of the bank and thus whether the bank experiences scale economies or scale diseconomies. The average *optimal* price cost margin for banks with total assets more than $654,000,000 (95th percentile) is around 7.3%, compared to the 5.8% found earlier. However, on average the *actual* price cost margin for these banks is 18.8%.[47] Interestingly, these high rents cannot be explained by cost advantages, as these banks are more likely to operate beyond the

[44] A disadvantage is of course the fact that this is a generated regressor, which requires us to be careful when discussing its significance.

[45] Since the main interest of the chapter is to examine whether the marginal effect of competition on innovation depends on the scale economies of firms, we follow Brambor et al. (2006) and use the standard deviation of the marginal effect to examine the significance of the conditional marginal effect as specified in column (iii). The conditional marginal effects of competition on the technology gap evaluated at the mean, median, 10th, 90th and 99th percentile are significant at the 1% level. For columns (iv) and (v), the coefficients of the interaction terms and the group dummy are jointly significant at the 1% level, but not individually. Furthermore, in column (v), the coefficients on competition and competition squared are jointly significant at the 1% level, but not individually.

[46] The instruments used in estimating the specification in column (*iii*) are jointly significant at the 1% level in the first stage regressions, with the competition variable, competition squared, firm size, the interaction between firm size and competition, and the interaction between firm size and competition squared as dependent variables. The F-values are 167.87, 102.35, 39.51, 40.55 and 23.32, respectively. Instruments from period $t-3$ and $t-4$ are used, since there is evidence of first-order autocorrelation in the residuals in levels. The instruments are relevant and valid according to first-stage F-tests and the Hansen test, respectively.

[47] Furthermore, the average annual *change* in the price cost margin is positive (0.7%) for these banks. Also interesting is the fact that 95% of the observations from the preferred specification in column (*iii*) in Table 3.2 have price cost margins higher than the optimal price cost margin associated with each observation. This means that banks are operating mostly on the left of the optimal point that enhances innovation in the inverted-U relationship between competition and innovation. However, we also allowed for a 5% deviation of the optimal price cost margin, since none of the banks operate exactly on the optimal price cost margin that enhances innovation. In this case, 88.7% of the observations have price cost margins below the optimal price cost margin minus the 5% deviation from this optimal point. Only 8.1% of the observations deviated less than 5% from the optimal price cost margin.

MES. The coefficients on the interaction terms indicate that the innovation behavior of larger banks is more responsive to changes in competition. Hence, the innovation incentives of large banks are diminishing if these banks continue to earn higher rents in the future.

In column (*iv*), we examine this interaction effect by comparing two groups, namely large and small banks.[48] Since the scale measure in the specification in column (iv) is a dummy variable, the coefficients on the interaction terms clearly capture the difference in the slopes of the inverted-U for banks operating below and above average scale economies (measured as total assets). The result is depicted in Figure 3.4. The optimal price cost margin for both groups is slightly positively. For the most part, large banks manage to have better technology sets. However, as competition decreases, and we move from the middle to the left in Figure 3.4, the technology gap drops at a much quicker pace for these large banks. Put differently, if consolidation proceeds, and competition decreases further, overall innovation (reflected by changes in the technology gap) by U.S. banks is expected to drop sharply, as these large banks dominate the market. Put simply, these large banks have indeed become 'too big to innovate'. Evaluated at the average price cost margin of 17.9%, the marginal effect of competition on the technology gap for banks with below and above average total assets is 11 and 7 percentage points, respectively. Hence, an increase in competition leads to relatively more innovation by smaller banks that operate near the average price cost margin.[49]

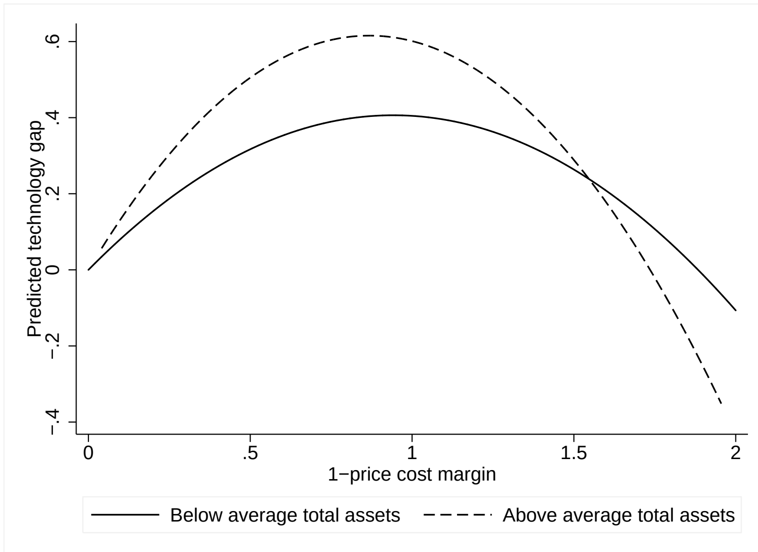
Column (*v*) shows the results when we employ a dummy (S) based on scale economies estimated using annual cost frontiers.[50] The dummy vari-

[48]For the specification in column (iv), the Hansen test indicates that the instrument set including lags from period t-3 and t-4 are valid based on a 1% significance level (which we apply considering the size of our dataset).

[49]The optimal price cost margin for banks with below and above average total assets is 5.8% and 13.2%, respectively. Smaller banks with a price cost margin near the average are operating relatively far away from the optimal point. Hence, the escape competition effect for this group is larger. We also evaluated the marginal effect by using one standard deviation below and above the average price cost margin. In both cases, the effect of competition on the technology gap is stronger for the larger banks. This indicates that these banks are more responsive to changes in competition.

[50]Lags from period t-3 are used as instruments for all endogenous regressors due to evidence of serial correlation in the residuals. For the dummy variable we use lags of the continuous scale economies variable as instruments. Lags from period t-4 are only used for the competition variable and the continuous scale economies variable. The inclusion of lags from period t-4 for the squared terms and interaction terms leads to a rejection of the null hypothesis that the instruments are valid (Hansen test).

Figure 3.4: The competition-innovation relationship



able separates groups with below and above average scale diseconomies. The results based on this alternative measure of scale economies also imply that the innovation behavior of banks with larger scale diseconomies is more responsive to changes in competition. Further decreases in competition lead to a relatively large drop in innovations by large banks. By evaluating the marginal effect of competition on technology gaps for both groups at the average price cost margin of 17.9%, we conclude that banks with below average scale diseconomies benefit more from an increase in competition. If the price cost margin increases by 1 percentage point, the technology gap increases by approximately 19.2 percentage points for banks with less scale diseconomies, while banks with larger scale diseconomies experience only an increase of 5.6 percentage points.[51]

Summing up, we have shown that large banks earn considerable rents, that they have become 'too big to innovate', and that the effect on innovation is sizeable. Added to the ongoing consolidation in U.S. banking, these results suggest that policies aimed at increasing competition may have an important positive externality: in addition to the usual downward pressure on prices,

[51]We also evaluated the marginal effect by using one standard deviation below and above the average price cost margin. The effect of competition on the technology gap is stronger for banks with larger scale diseconomies in both cases. This indicates that these banks are more responsive to changes in competition.

increased competition may, ironically, result in further cost reductions, as large banks in particular try to escape competition by innovating.

3.5 Conclusion

This study has examined the effect of scale (dis)economies on the competition-innovation relationship in U.S. banking. Theoretical endogenous growth models have ignored this effect, as they rely on the premise that unit costs are independent from output levels. Such models are less appropriate for investigating the effect of competition on innovation in sectors where average costs are not constant.

We have extended the theoretical model of Aghion and Griffith (2005) to allow for a U-shaped average cost curve where firms may operate below, on or beyond the MES. The latter constitutes a novelty of this chapter as it allows us to derive conditions under which the magnitude of the Schumpeterian effect and escape competition effect differs between firms. Firms that operate below the MES have a higher bonus when they innovate since they experience more potential for growth. The (process) innovation lowers their average costs, and thereby increases the expected rents from innovating. Firms that operate beyond the MES have less growth potential since firm growth may increase their average costs.

In addition we have introduced a novel way to measure process innovation in banking. The measure of innovation that we utilize focuses on banks' ability to minimize costs through innovations. Innovations improve the technology sets of banks and narrow the technological distance between the technology applied in a bank and the best (potential) available technology. Relying on recent contributions to the estimation of meta cost frontiers, we measure these technology gaps.

Subsequently, we have tested the theoretical implications of our model on a rich data set of U.S. banks. Our main aim was to find whether large U.S. banks have become 'too big to innovate'. We find that most banks in the U.S. start to operate beyond the Minimum Efficient Scale and experience scale diseconomies as the sector consolidates and average bank sizes increase. The upward trend in the average price cost margins during the same period implies that the degree of competition in the banking sector has declined. Our results provide support of an inverted-U relationship between competition and innovation in U.S. banking. This finding is robust over several different model specifications and consistent with the theoretical and empirical work of Aghion et al. (2005b) and Bos et al. (2009).

Further analysis has revealed the effect of scale (dis)economies on the nature of the competition-innovation relationship in U.S. banking. Large banks are shown to earn considerable rents. Important, the inverted-U relationship between competition and innovation becomes steeper as bank size increases. Our findings have important implications for competition policy. As banks, on average, are becoming larger over time, the increased responsiveness in terms of innovation should be taken into account when implementing competition policies. Any further decreases in competition will prove to be highly detrimental to innovation, due to the more than proportional reduction in innovation by large banks that have become 'too big to innovate'. In contrast, the added bonus of any policy that effectively stimulates competition is a boost in innovation, due to its effect on large banks.

Chapter 4

The Effect of Innovation Behavior and Performance on Competition

4.1 Introduction

Since the seminal work of Schumpeter (1942), a large literature on the relationship between competition and innovation emerged in the fields of Industrial Organization and Endogenous Growth. One of the main difficulties in these papers is to examine the causality between competition and innovation. A majority of the papers focuses on the effect of competition on innovation and conclude that competition is an important determinant of the innovation incentives of firms.[1] Aghion et al. (2005b) provide an important contribution to this literature by integrating positive and negative effects of competition on innovation into one framework to explain the existence of an inverted-U relationship between competition and innovation. Relatively less attention is paid to the causality running from innovation to competition. Furthermore, the empirical literature on the effect of innovation on competition is still inconclusive.[2] Farber (1981) finds a positive

[1] See Kamien and Schwartz (1982), Symeonidis (1996) and Ahn (2002) for a review of this literature.

[2] Specifically, most papers examine the effect of innovation on market structure and (implicitly) assume that more concentrated markets are less competitive. However, the relationship between concentration and competition is unclear. Ivaldi et al. (2003) argue that tacit collusion is more easy to sustain in markets with a few firms. Nevertheless, concentrated markets may be very competitive if firms engage in Bertrand competition with homogenous products.

effect of the employment of scientists and engineers on seller concentration. Lunn (1986) distinguishes between process and product patenting and finds that process patenting leads to more concentrated markets. His empirical results suggest that product innovations and concentration are not related. However, negative effects have also been found in the literature. Mukhopadhyay (1985) finds lower concentration ratios in R&D-intensive industries and Geroski and Pomroy (1990) find that innovation counts lead to less concentrated industries. Several theoretical contributions hint at a positive effect of innovations on market concentration. In the model of Jovanovic and MacDonald (1994), exogenous technological developments elicit the exit of firms where the number of firms rises initially, but declines due to the inability of several firms to adopt the new technology successfully. Klepper and Simons (2000) use a model of industry evolution and examine the technology choices of firms and their survival chances. In their model, technological change and the early adoption of new technologies plays an important role in firm survival and the consolidation of markets.[3] Despite the empirical and theoretical contributions on this subject area, there is no consensus on the effect of innovation on competition. An important limitation in the extant literature is the focus on innovation output or innovation input in isolation. These studies are unable to examine within one framework whether innovation output has a different effect on competition than innovation input. For example, firms may become more dominant in a market after surpassing other competitors due to successful innovations. However, firms may also behave more competitively if they expect to gain a technological advantage over other firms in the future by increasing their efforts into R&D. Examining the effects of innovation input and innovation output in one framework is important since the exclusion of one of these variables may not only lead to omitted variable bias, but makes it impossible to examine the relative importance of innovation input and output for the degree of competition. Omitted variable bias makes it more difficult to assess whether innovation output and input affect competition positively or negatively.

The purpose of this chapter is to examine within one framework whether innovation output and input have different effects on competition. Furthermore the relative importance of these determinants of competition are investigated. Two-yearly Dutch manufacturing data from the Community Innovation Survey (CIS) and Production Statistics (PS) over the period

[3]Both Jovanovic and MacDonald (1994) and Klepper and Simons (2000) use their models to explain developments in the U.S. tire industry. This industry experienced a severe shake-out since technological change affected the survival of firms. Moreover, the U.S. tire industry evolved towards a tight oligopoly.

1994-2004 are used.

Examining the causal effect of innovation on competition is also of interest for policy purposes. Investigating the effects of innovation on competition is important since innovation policies with respect to patent rights and R&D subsidies can be used to influence competition. Patents provide firms the right to exclude competitors and therefore has anti-competitive effects on competition. Hence, altering patent policies by implementing more or less stringent patent laws affects the market power of a firm with a patent. R&D subsidies influences firms' their innovation efforts and their expectation of their competitive position in the future. In turn, this expectation alters their current competition strategies. Examining the effects of innovation on competition is also important for competition policies used to foster innovation in industries. Specifically, feedback effects from innovation on competition may amplify or reduce the competition stimuli from policies.

This study contributes to the literature in two ways. First, the effect of innovation output and input on competition is examined in one framework. This approach allows the identification of the relative importance of innovation output and input for competition. Second, new innovation measures from the Dutch Community Innovation Survey are used in the context of this study. The percentage sales from new or improved products is used as a measure of innovation output and captures the successfulness of innovations. The cumulative innovation intensity based on the accumulation of total innovation expenditures divided by total sales is used as a measure of innovation input (accumulated experience) and captures a large variety of innovation inputs in addition to R&D expenditures. These measures are not used in the extant literature on the effects of innovation on competition.

The rest of the chapter is structured as follows. Section 4.2 describes the model used to explain tacit collusion within sectors and the effect of innovation output and input on tacit collusion. Section 4.3 overviews the data and methodology. Section 4.4 provides an overview of the empirical evidence on the effects of innovation input and output on competition. Section 4.5 concludes.

4.2 Theoretical framework

In this section, a tacit collusion model is used to examine how innovation behavior and performance affect competition. Competition is modeled in this chapter by the sustainability of tacit collusion. Tacit collusion refers to the conduct of firms that leads to collusive outcomes in a noncooperative manner (Tirole, 1988). The basic model presented here is based on Tirole (1988) and Ivaldi et al. (2003). However, the focus in this section is also on the innovation output instead of on the probability to innovate only as in Ivaldi et al. (2003). The basic model of Tirole (1988) and Ivaldi et al. (2003) is extended by modeling the probability to innovate as a function of the stock of knowledge as an innovation input.[4]

There are n firms that produce on the same market with the same marginal costs over an infinite horizon ($T = +\infty$). The firms are assumed to engage in Bertrand competition. Firm i earns profit $\pi_{it}(p_{1t}, ..., p_{nt})$ at each date t, where $i \in \{1, ..., n\}$, $t \in \{0, ..., T\}$ and p_{it} are the prices that firms charge. Each firm maximizes the present value of its expected profits:

$$E(V) = \sum_{t=0}^{T} \delta^t E(\pi_{it}(p_{1t}, ..., p_{nt})), \tag{4.1}$$

where δ is the discount factor.[5] Firms consider the following grim trigger strategies:

$$p_{it}(H_t) = \begin{cases} p_c & \text{if } H_t = (p_{c10}, ..., p_{cn0}; ...; p_{c1t-1}, ..., p_{cnt-1}) \\ p = mc & \text{otherwise,} \end{cases} \tag{4.2}$$

where H_t contains the history of the price setting behavior of firms. All firms charge the collusive price p_c in period $t = 0$. This collusive price is charged in period t if all firms charged the collusive price in the previous periods. If one firm deviates from the collusive price in the previous period, firms set prices at their marginal costs mc forever. We assume that innovations may occur after the initial period, that one of the incumbent firms innovates with probability θ and earns expected profits V_I. If one incumbent innovates, it captures all the rents in the market and other firms earn zero profits. The probability to innovate for firm i is an increasing concave function of the stock of knowledge as an innovation input $\rho_i = f(KS_i)$.[6] If firms commit themselves to charge the collusive price, the expected discounted profits are:

$$E(V_C) = \frac{\Pi_C}{n} + \sum_{t=1}^{T} \delta^t (1-\theta)^{t-1} \left(\frac{\theta}{n} V_I + (1-\theta) \frac{\Pi_C}{n} \right). \tag{4.3}$$

[4]Furthermore, the case with n firms is shown in this study instead of two firms as in Tirole (1988) and Ivaldi et al. (2003).

[5]The discount factor equals $1/(1+r)$, where r is the discount rate. Time is discrete in this model instead of continuous.

[6]The term ρ_i represents the probability that firm i innovates, while the term θ represents the probability that one of the n firms in the market innovates.

The first term on the right-hand side shows that all firms earn a share of the total collusive profits in the market Π_C since it is assumed that innovations may occur after the initial period. The second term shows that firms discount their expected profits based on their probability to survive in future periods. Rewriting equation (4.3) gives:

$$E(V_C) = \frac{\Pi_C}{n(1-\delta(1-\theta))} + \frac{\delta\theta V_I}{n(1-\delta(1-\theta)).} \tag{4.4}$$

If a firms decides to defect by not adopting the collusive price, it captures the whole market in the short-run and earns Π_C, but elicits a price war where firms price at their marginal costs. Hence, profits will be zero if no innovations occur. If a firm manages to innovate successfully, it will earn expected rents V_I. The expected discounted profits for a deviant firm are given by:

$$E(V_D) = \Pi_C + \sum_{t=1}^{T} \delta^t (1-\theta)^{t-1} \left(\frac{\theta}{n} V_I\right), \tag{4.5}$$

where firms discount only the expected profits from innovation in future periods since they earn zero profits if the other firm innovates or if no innovation occurs at all (all firms price at marginal costs). Rewriting equation (4.5) gives:

$$E(V_D) = \Pi_C + \left(\frac{\delta\theta V_I}{n(1-\delta(1-\theta))}\right). \tag{4.6}$$

Firms in the market collude if the expected discounted profits from charging the collusive price is higher than the expected discounted profits from pricing at marginal costs (i.e. $E(V_C \geq V_D)$). Hence, collusion occurs if the discount factor is higher than a threshold at which collusion is sustainable, which in turn depends on the probability that an innovation occurs in the market. Rewriting the condition $E(V_C) \geq E(V_D)$ to the minimum value of the discount factor required to sustain collusion gives:

$$\delta \geq \delta^*(\theta, n) = \frac{n-1}{n(1-\theta)}, \tag{4.7}$$

where δ^* represents the threshold value of the discount factor that sustains collusion. Firms collude if the discount factor exceeds a critical value. Differentiating equation (4.7) with respect to the probability to innovate θ gives:

$$\frac{\partial\delta}{\partial\theta} = \frac{n^2 - n}{(n(1-\theta))^2} > 0. \tag{4.8}$$

Equation (4.8) shows that it becomes more difficult to collude if the probability of innovation increases. The intuition behind this result is that firms obtain a dominant position in the market after they innovate. This possibility makes the expected profit from collusive behavior less valuable. Furthermore, innovations lower the strength of the other firms to retaliate against non-collusive behavior. Since the probability to innovate is an increasing function of the innovation input (KS), higher levels of innovation input increase the probability to innovate and therefore hinder collusive outcomes. However, while competitive behavior becomes more likely when the innovation input increases, the innovation output has anti-competitive effects. The innovating firm captures the market, experiences an increase in its market power and sales due to the innovation S_I and earns profits V_I. Hence, the innovating firm experiences a higher price cost margin $PCM_I = V_I/S_I$ than non-innovating firms $PCM = 0$.[7]

The model presented here is just a simplified case where the probability to innovate depends on the level of the stock of knowledge. The model shows that competition increases if the probability to innovate increases (e.g. due to increases in the innovation input). However, while innovation input positively affects competition, an innovative firm becomes more dominant in the market after a successful innovation and experiences higher sales and price cost margins. Thus innovation output reduces competition in a market.

4.3 Data and Methodology

4.3.1 Data

Dutch innovation statistics at the firm-level are obtained from the Community Innovation Survey (CIS) and merged with financial information of enterprises from the Production Survey (PS). The data are collected by the *Centraal Bureau voor de Statistiek*. The bi-annual CIS data stem from five survey waves, namely CIS 2 (1994-1996), CIS 2.5 (1996-1998), CIS 3 (1998-2000), CIS 3.5 (2000-2002) and CIS 4 (2002-2004). Firms in all sectors with 10 or more employees are included in the CIS data. The PS and CIS data

[7] An extreme case is shown in this section since non-innovators earn zero profits. However, the conclusion that innovating firms earn higher price cost margins also holds if we assume that non-innovators earn nonzero profits since innovating firms capture most of the market.

are based on a census and a stratified random sample. The census data contain the population of firms with 50 employees or more and the stratified random sample is based on firms with less than 50 employees. The stratum variables are the economic activity and firm size as measured by the number of employees.[8] Firms that are included in one survey only are excluded from the sample.[9]

The CIS questionnaires consist of two parts. In the first part, firms are asked to provide general information concerning their economic activity, sales, number of employees, exports etc. The second part contains questions related to innovation activities (R&D, turnover from innovations as percentage of total sales, other innovation input expenditures etc.). Firms are asked to answer the questions in the second part if they affirmed one of the three questions regarding: 1) whether firms developed new or strongly improved products; 2) whether firms have put new or strongly improved production processes into use; 3) whether the firm has ongoing or abandoned innovation activities. Firms are defined as innovators if they affirm at least one of these three questions.

4.3.2 Measuring competition

In this study, competition is measured by using the price cost margin as in Aghion et al. (2005b).[10] The price cost margin is calculated by dividing the total sales minus the cost of sales, labor expenses and energy costs by total sales:

$$C_{it} = (\frac{S_{it} - TVC_{it}}{S_{it}}), \tag{4.9}$$

where C_{it} is the competition variable, S_{it} total sales and TVC_{it} represents the total variable costs. A firm-level competition measure is used in this chapter instead of an industry-level based measure, since industries are relatively broadly defined in the dataset.[11] Hence, it is assumed that all changes in competition are reflected in the price cost margins of firms.[12] The price

[8]Regressions based on firms with only more than 50 employees are also performed as a robustness check. The exclusion of firms with less than 50 employees has a negligible effect on the parameter estimates.

[9]Furthermore, we assume that attrition of the panel data occurs exogenously.

[10]Specifically, Aghion et al. (2005b) use one minus the price cost margin.

[11]The classification is based on the 3-digit SBI. SBI is the Dutch standard industrial classification and represents the economic activity of a firm.

[12]An important assumption in this thesis is that less competition is positively related with higher price cost margins, while more competition leads to lower price cost margins.

cost margin has several advantages over conventional measures competition such as the Herfindahl-Hirschman Index (HHI) and concentration ratios. Unlike the price cost margin, the HHI and concentration ratios require a precise definition of the relevant geographical or product boundaries. The price cost margin is restricted between the values -1 and 1.[13]

4.3.3 Measuring innovation output: percentage sales from innovations

The sales from new or improved products divided by total sales is utilized as a measure of innovation output.[14] Many studies on innovation use patents as an innovation output measure. The limitations of patents are elaborately discussed in the literature.[15] For example, not all innovations are patented, some patents are never translated into commercially viable products and it may be difficult to capture the economic value of patents. The main advantage of using the percentage sales from new innovations as an innovation output indicator is that it captures innovation directly by measuring the introduction of a new product or service and the success of the innovations. A limitation of this measure is that some firms may provide rough estimates of the shares in sales of innovative products. Consequently, this regressor may be subject to measurement error. Another limitation is that the measure may be influenced by the life cycle of a product. Firms where more than 50% of their total sales consists of sales from innovations new to the firm are excluded from the sample.[16] Furthermore, firms with only process innovations are also excluded from the sample. The percentage sales from innovations does not capture the output of process innovators since these firms do not have sales from innovations.[17]

[13] In total, 8 observations are excluded due to this restriction.

[14] As in Brouwer et al. (2008), the main analysis is restricted to products new to the firm instead of basing the innovation output measure on products new to the market. They argue that a measure based on products new to the market may suffer from problems associated with the interpretation of firms of their scope of the relevant market. As a consequence, firms that are more oriented at home markets may overestimate their innovative efforts.

[15] See Kamien and Schwartz (1982), Acs and Audretsch (1987b) and Geroski (1990).

[16] This restriction is also used by Raymond et al. (2010) and leads to the exclusion of 806 observations.

[17] In total, around 1300 observations are excluded. The inclusion of process-innovators has a negligible effect on the parameter estimates.

4.3.4 Measuring innovation input: cumulative innovation expenditures divided by sales

While many papers do not account for learning effects in patent races, it is assumed in this study that a firm's probability to innovate (θ/n) does not depend on the flow of innovation expenditures but on the accumulated innovation experience. Doraszelski (2003) shows R&D efforts in the past affect the probability to win the R&D race positively and that firms have incentives to reduce R&D expenditures if their knowledge stock increases. Hence, a measure based on the cumulative total innovation expenditures is used to represent the accumulated innovation experience. An advantage of the CIS waves is that they contain more information on innovation input than R&D expenditures. An important limitation of using R&D expenditures instead of total innovation expenditures is that R&D expenditures may understate the research activity of small firms. Furthermore, R&D expenditures do not capture other innovation input expenditures. Total innovation expenditures cover more innovation inputs than R&D expenditures and consist of (total) R&D expenditures (internal and external), the purchase of rights and licenses to use external technology, and the purchase of advanced machinery and computer hardware devoted to the implementation of product and process innovations.[18] Hence, a measure of accumulated innovation experience is constructed based on the total innovation expenditures:

$$I_{it} = \frac{\sum\limits_{t=0}^{t} TIE_{it}}{S_{it}}, \tag{4.10}$$

where I_{it} is the innovation input variable, TIE_{it} is the total innovation expenditure and S_{it} are total sales. Accumulated innovation experience is normalized by sales since accumulated expenditures in levels may understate the knowledge intensity of small firms. A limitation of this measure is that some firms may provide rough estimates of total innovation expenditures. Consequently, this regressor may be subject to measurement error. Another limitation is that this measure may negatively affect the response rate of questionnaires. Observations are excluded if values of the total expenditures divided by sales are smaller than 0 or larger than 1.[19]

[18] The pre-CIS4 data also includes additional components in the total innovation expenditures. However, these components are not used in this study since some components are not included as continuous measures in CIS4.

[19] This restriction leads to the exclusion of 185 observations.

4.3.5 Control variables

Two dummy variables are included as control variables to account for the influence of cooperative innovation projects on competition. The two cooperation dummy variables indicate whether firms cooperate with competitors ($COOPCOMP$) or other institutions ($COOPOTHER$). The reference group are firms that do not cooperate. In general, cooperating firms are engaged in higher levels of innovative activities (Tether, 2002; Belderbos et al., 2004). In turn, higher levels of innovative activities increase the probability to innovate and may result in more competitive behavior. However, cooperation with competitors may have an ambiguous effect on competition. Firms may behave more competitively by increasing the probability to innovate via cooperative R&D activities. However, Wu and Wei (1998) argue that firms may use R&D cooperation as a device to achieve tacit collusion. This outcome is more likely if the efficiency gains due to cooperative R&D activities are small.[20]

A dummy variable is also included to control for whether firms introduce product innovations only ($PRODONLY$) or both product and process innovations. This dummy variable has an ambiguous effect on tacit collusion according to the theoretical literature. Product innovations can be vertically or horizontally related to existing products, where a vertical relation refers to quality improvements in existing products and a horizontal relation refers to new products that expand the variety in consumption or specialization in production. Vertical product differentiation hinders tacit collusion since the high quality firm is more tempted to lower prices in an attempt to enjoy a higher profit margin (Ivaldi et al., 2003). Hence, this firm gains more from stealing additional customers from its rivals. An increased degree of horizontal product differentiation may have ambiguous effects on tacit collusion.[21] It becomes more difficult for firms to capture other firms' (sub)markets by charging competitive prices, but increased product differentiation also limits the severity of retaliation if firms deviate from collusive pricing. Process innovations also have an ambiguous effect on tacit collusion. If some firms produce relatively more (or better) process innovations, cost asymmetries may hinder (tacit) collusive behavior. Cost asymmetries between firms impede tacit collusion since firms with lower marginal costs gain relatively more by pricing below the other firms' marginal costs due to

[20]Wu and Wei (1998) find no support that cooperative R&D leads to tacit collusion among cooperating firms.

[21]Deneckere (1983), Chang (1991), Rothschild (1992), Albk and Lambertini (1998) examine the effect of product differentiation on the ability to collude.

the weak retaliatory capabilities of competitors with higher marginal costs. However, process innovations may also affect scale economies which in turn may impose entry barriers in a market. High entry barriers facilitate collusion by shielding colluding incumbents from possible competition by entrants (Feuerstein, 2005).

Firm size is measured as the total sales (S) and used as a control variable since the size of firms may also affect the degree of competition. For example, larger firms may gain a competitive advantage if they are able to exploit economies of scale in production (Gaba et al., 2002).

A dummy variable is used to indicate whether firms are continuously or occasionally innovating within each CIS questionnaire ($INNOFREQ$). Its effect on the competition variable is ambiguous. A firm that innovates continuously may have a higher probability to innovate and hence behave more competitively according to the model presented in section 4.2. However, continuously innovating firms may have a better competitive position compared to firms that innovate occasionally.

Industry effects and unobserved firm-specific fixed effects are included in the specifications. The industry effects capture industry-specific heterogeneity (e.g. technology classes, entry barriers etc.) and the firm-specific fixed effects may capture firm heterogeneity such as firm-specific technological opportunities and approachability conditions related to competition and innovation. Technological opportunity conditions reflects factors that affect the likelihood that an innovation occurs given the amount of investment in R&D and other innovation inputs, while appropriability conditions reflect the possibility of protecting innovations from imitation (Baptista and Swann, 1998).

Table 4.1 provides the descriptive statistics of the innovation output, input and control variables. The means of the main variables of interest, namely the price cost margin, innovation output and input are 24.2%, 18.8% and 6.2%, respectively. Most of the firms are not cooperating on innovation activities according to these descriptive statistics. Only 14.2% of the firms cooperate with competitors and 42.6% cooperate with other institutions. More than half of the firms performs innovation activities in isolation, namely 56%. Around 34% of the firms in the sample produce product innovations only and 66% produce both product and process innovations. On average, firms have total sales around 63,000,000 EUR. Around 72.1% of the firms in the sample are continuously innovating, while 27.9% innovates occasionally.

Table 4.1: Descriptive statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
Price cost margin	0.242	0.125	-0.980	0.886
% Sales from innovations	0.188	0.134	0	0.5
Cumulative innovation expenditures / Sales	0.062	0.100	0	0.994
Cooperation with competitors	0.142	0.350	0	1
Cooperation with other institutions	0.426	0.458	0	1
Product innovations only	0.340	0.474	0	1
Sales per 1,000,000 EUR	62.811	236.897	0.097	4158.635
% of continuously innovating firms	0.721	0.449	0	1

The descriptive statistics are based on the sample of specification 1 in Table 4.2 (4,151 observations).

4.3.6 Methodology

The empirical specifications are based on the following model:[22]

$$C_{it} = O_{it}\beta + I_{it}\delta + \gamma' Z_{it} + a_i + \varepsilon_{it} \tag{4.11}$$

where C_{it} is the competition variable, O_{it} is the innovation output variable, I_{it} represents the innovation input variable, Z_{it} is a vector of control variables, a_i is the unobserved heterogeneity assumed to be constant over time, and the error term ε_{it} is a white noise process.

First, a basic fixed effects model without control variables is estimated (specification 1). Second, control variables are included to examine the effect of cooperation with other institutions, the type of innovations, the frequency at which innovations take place and firm size on competition (specification 2). Third, the chapter proceeds with an important issue in studies on the relationship between competition and innovation, namely simultaneity bias. While competition affects innovation by altering innovation incentives, the theoretical model presented in section 4.2 shows that increases in innovation inputs may hinder collusive behavior by increasing the probability that innovations occur. Furthermore, the model shows that innovation output may lead to a more dominant position in the market after successful innovations. Since the interest of this study lies in examining the causal effect of innovation input and output on competition, instruments are used for the

[22]Although process innovators are excluded from the sample, the same model can be used if process innovation output and input are available. In general, it is argued in this chapter that more innovation input increase competition, irrespective whether it concerns a product or process innovation. Furthermore, firms may become more dominant after they implement process innovation successfully.

endogenous regressors. Several instruments are used in the estimation procedures. The cumulative amount of researchers, assistants and other personnel related to R&D activities in full-time equivalent are used as instruments for the cumulative total innovation expenditures divided by total sales. It is assumed that the amount of R&D related personnel is correlated with total innovation expenditures expenditures since more R&D related personnel may require more expenditures on R&D equipment. Furthermore, it is assumed that this instrument is uncorrelated with the error term in equation (4.11) and that competition is not a direct function of R&D related personnel.[23] A dummy variable is used as an instrument for the innovation output variable and captures whether the firm engaged in activities for the market introduction of the new or significantly improved goods and services.[24] For example, firms with intensive advertising activities for new products may attain more sales from innovations. In specification 3, only innovation input is treated as an endogenous regressor. In specification 4, both innovation output and innovation input are treated as endogenous regressors.

Several robustness checks are performed in this study. First, equation (4.11) is estimated by using the First-Differences estimator (specification 5). Large differences between the results based on fixed effects and first differences may indicate problems associated with the stationarity of the variables, biased estimates due to model misspecification or measurement error. Second, interaction terms are included in equation (4.11) to examine whether innovation output and input have different effects on competition in the manufacturing and services sector (specification 6).[25] Sirilli and Evangelista (1998) point out that the risk to be imitated by competitors is a more important hampering factor to innovation in manufacturing than in services. They conclude that appropriability conditions are more important determinants of technological change in manufacturing than in services. Hence, introducing and selling new innovative products may have a stronger effect on profitability and the degree of competition in the services sector. Sirilli and Evangelista (1998) also argue that innovating is more costly in manufacturing by examining total innovation expenditures per employee.[26] The

[23] The quality of R&D personnel is an important determinant of successful innovation and the decisions about (global) R&D investments (Dearing, 2007). Hence, it is assumed in this chapter that the quality of R&D related personnel is an important innovation input that may affect competition directly instead of the amount of R&D related personnel.

[24] The activities include market research and launch advertising.

[25] The value 1 of the dummy variable indicates manufacturing firms and the value 0 represents firms in the services sector. This dummy variable is interacted with the innovation output and innovation input measure separately.

[26] Sirilli and Evangelista (1998) investigate the innovation behavior of firms in Italy.

same result is found when the total innovation expenditures are examined with Dutch CIS data. Manufacturing firms spent on average 6,000 euro per employee, while services firm spent on average 4,500 euro per employee over the period 1994-2004.[27] Therefore, innovation input as measured by total innovation expenditures may have a stronger effect on competition in manufacturing than in services. Third, a different measure of accumulated innovation experience based on Xu and Zhang (2004) is utilized as a robustness check (specification 7). This measure is constructed by accumulating the innovation expenditure intensities over time.[28]

4.4 Results

4.4.1 Regression results

In this section, specification 1-4 are used to examine whether innovation output leads to less competition and whether innovation input leads to more competition. In specification 1 (fixed effects without controls), both the innovation output and input variable are individually significant at the 1% level.[29] A one percentage point increase in innovation output increases the price cost margin by 0.064 percentage point. The proxy for accumulated innovation experience is also significant at the 1% level and has a negative effect on the price cost margin. An increase in the innovation input variable by one percentage point increases the price cost margin by 0.175 percentage point.[30]

Specification 2 is an extension of specification 1 by adding several control variables to examine the effect of cooperation with other institutions, the

[27] The means differ significantly from each other at the 5% level. The spread in innovation expenditures per employee is also larger in the manufacturing sector.

[28] The innovation intensity is based on total sales instead of total assets as in Xu and Zhang (2004). Total assets are not included in the CIS and PS data. Furthermore, weights on past intensities as in Xu and Zhang (2004) are not used. The weights depend on the assumptions concerning the depreciation rate of knowledge.

[29] Regressions based on the innovation output and innovation input only show similar results. Omitted variable bias based on the exclusion of innovation output or innovation input does not seem to be a major problem in this study. The coefficient of the innovation output variable is somewhat higher (0.080) if the innovation input variable is excluded. The correlation coefficient between the innovation output and innovation input measure is 0.1128. The relationship between innovation output and innovation input is examined more elaborately in chapter 5.

[30] The coefficient of the innovation input and output variable differ significantly from each other at the 1% level.

Table 4.2: Estimation results: main specifications

Specification	(1) *Fixed effects*	(2) *Fixed effects*	(3) *2-step GMM*	(4) *2-step GMM*
O_{it}	0.064***	0.069***	0.061***	0.690
	(0.012)	(0.012)	(0.014)	(0.698)
I_{it}	-0.175***	-0.174***	-0.227**	-0.043
	(0.036)	(0.037)	(0.097)	(0.229)
$COOPCOMP_{it}$		0.006	0.009*	0.004
		(0.005)	(0.005)	(0.009)
$COOPOTHER_{it}$		-0.005	-0.006	-0.002
		(0.004)	(0.004)	(0.006)
$PRODONLY_{it}$		-0.002	-0.001	0.029
		(0.004)	(0.004)	(0.035)
S_{it}		0.0001***	0.0001**	0.0001
		(0.00002)	(0 00003)	(0.00007)
$INNOFREQ_{it}$		-0.004	-0.004	-0.021
		(0.004)	(0.004)	(0.019)
Observations	4,151	4,057	4,057	4,057
Hansen J statistic			1.126	0.406
			(0.570)	(0.816)

The dependent variable is the competition variable (price cost margin). Standard errors (between parentheses) are robust against heteroskedasticity and serial correlation. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01. The chi-squared statistic and p-value (between brackets) are reported for the Hansen test. Only instruments for the innovation input variable are used in specification 3. In specification 4, instruments are used for both the innovation output and innovation input variable.

type of innovations, the frequency at which innovations take place and firm size on competition.[31] The parameter estimates are not severely influenced by the inclusion of control variables. The coefficients of the innovation output and input variable are 0.069 and 0.174, respectively. Most of the control variables are insignificant. Consistent with the results of Wu and Wei (1998), the findings in this chapter do no support the notion that cooperative innovation activities lead to tacit collusion among cooperating firms. Cooperation with other institutions, the type of innovations and the frequency at which innovations take place are also not related to competition. Only the total sales is significantly positively related to competition at the 1% level. However, it is not clear whether this relationship is due to the possibility that larger firms behave less competitively or due the relationship between price cost margins and sales by construction. Nevertheless, the results based on specification 1 and 2 suggest that more successful innovations (output) are related to less competition, while higher levels of innovation input (measured as the accumulated innovation experience) is related to more competition. However, the parameter estimates of the variables of interest (innovation output and input) are subject to simultaneity bias in specification 1 and 2.

In specification 3, only the innovation input variable is treated as an endogenous regressor and instruments based on the amount of researchers, assistants and other R&D related personnel are used. The instruments are exogenous and relevant according to the Hansen test and an F-test to examine the joint significance in the first-stage regression, respectively.[32] The innovation output measure is significant at the 1% level. Furthermore, the coefficient of the innovation output variable is quite similar to the coefficient in specification 1 and 2. Also the magnitude of the coefficient of the innovation input variable differs slightly from the estimated coefficients in specification 1 and 2. An increase of one percentage point in the innovation input variable decreases the price cost margin with 0.227 percentage point instead of approximately 0.175 percentage point in the first two specifications. An explanation for the less negative coefficient of the innovation input variable in specification 1 and 2 may be due to the effect of competition on innovation incentives. The causal effect of innovation on competition cannot be captured since no instruments are used in the first two specifications. Aghion et al. (2005b) examine the relationship between competition and innovation and argue that there is an inverted-U relationship between com-

[31] The Hausman specification test based on specification 2 indicates that a fixed effects model is preferred over a random effects model.

[32] The F-statistic is 17.05 is larger than the often used rule of thumb of 10.

petition and innovation. There is a dominant positive (escape-competition) effect of competition on innovation at lower levels of competition until competition reaches an optimal point that enhances innovation. Beyond this optimal degree of competition, a negative (Schumpeterian) effect dominates the positive effect. The differences between the results without instruments (specification 1 and 2) and with instruments (specification 3), indicate that there may be a dominating negative effect of competition on the innovation incentives of firms. The dominating negative effect of competition on innovation leads to a smaller positive relationship between innovation input and competition in specification 1 and 2.[33] As in specification 2, cooperation with other institutions, the type of innovations, the frequency at which innovations take place are not related to competition. The results only show weak evidence with respect to the relationship between cooperation with competitors and tacit collusion. The positive effect of cooperation with competitors on the price cost margin is only significant at the 10% level. The variable total sales is significantly positively related to the price cost margin.

In specification 4, both the innovation input and output variable are instrumented to capture their causal effect on competition. A dummy variable that indicates whether the firms engaged in marketing-related activities to promote new products or services is used as an instrument for the proportion of sales from new products or services (innovation output). As in specification 3, the instruments for the innovation input variable are based on the amount of researchers, assistants and other R&D related personnel. The instruments are exogenous according to the Hansen test. However, the instrument set for the innovation output variable is not relevant.[34] The irrelevance of the instruments seems to affect the precision of the estimates since both the innovation output and input variable are individually insignificant. However, the proportion of sales from new products or services and the proxy for accumulated experience are jointly significant at the 10% level. Thus, although the estimates seem to be heavily influenced by the weak instruments, specification 4 shows weak evidence that innovation output and innovation input affect competition jointly. Nevertheless, the direction of the effect is difficult to assess. The results also show that there is no relationship between cooperation with other institutions, the type of innovations, the frequency of innovations, firm size and competition.

[33]Specifically, it may lead to a less negative relationship between the innovation input and price cost margin.

[34]The F-statistics in the first-stage regressions for the innovation output and input

Table 4.3: Standardized coefficients

Specification in Table 4.2	(2)	(3)	(4)
O_{it}	0.036***	0.032***	0.358
I_{it}	-0.022***	-0.029**	-0.006

Table 4.3 shows the standardized parameter estimates based on specification 2, 3 and 4 in Table 4.2 to examine the relative contribution of the innovation input and output variable to competition.[35] Standardizing the coefficients of specification 2 shows that innovation output is relatively more important for competition since the absolute value of the standardized coefficient is higher than the standardized coefficient of the innovation input variable. However, the parameter estimates are not reliable due to simultaneity bias. Specification 3 shows the results with standardized coefficients and instruments for the endogenous innovation input variable. The findings in specification 3 support the outcomes based on the standardized coefficients of specification 2. The absolute value of the coefficient of the innovation output variable is higher than the standardized coefficient of the innovation input variable. Hence, innovation output is relatively more important for competition than innovation input. However, only the endogenous innovation input variable is instrumented in this specification. Both the innovation output and input variable are instrumented in specification 4 and all coefficients are standardized. Innovation output and innovation input are jointly significant at the 10% level, but not individually significant. While both variables affect competition jointly, it is difficult to assess the relative magnitude of variables since the results may be influenced by the weak instrument for innovation output.

In sum, the results based on specification 1-3 in Table 4.2 are consistent with the theoretical expectation that firms become more dominant after successful innovations. Furthermore, these outcomes are in line with the theoretical model in section 4.2 which predicts that increases in the probability to innovate (due to more innovation experience) hinder tacit collusion. The results based on standardized coefficients also suggest that innovation

variable are 1.69 and 13.67, respectively.

[35] All variables are transformed by subtracting the mean and dividing by the standard deviation. All variables have a mean of 0 and a standard deviation equal to 1 after the transformation of the variables. The standardized parameter estimates indicate the change in competition (in standard deviations) due to a change in the innovation input or output of one standard deviation. The standardization does not affect the statistical significance of the variables.

output is relatively more important for competition than innovation input. However, the outcomes based on specification 4 show that innovation output and input affect competition jointly, but provide no clear evidence on the direction and relative magnitudes of the effects.

4.4.2 Robustness checks

Table 4.4 contains the results of the robustness checks. Specification 5 is based on equation (4.11). However, the First-Differences estimator is utilized to eliminate the unobserved heterogeneity that is constant over time. The results based on specification 5 show that firms with higher sales from new products or services relative to their total sales have higher price cost margins. The innovation output variable is significant at the 1% level. A one percentage point increase in the percentage sales from new products leads to an increase in the price cost margin of 0.038 percentage point. The magnitude of this effect is comparable with the results based on specification 1-3 in table 4.2. The coefficient of the innovation input variable is negative and significant at the 5% level. A one percentage point increase in the innovation input variable leads to a decrease of 0.151 percentage point in the price cost margin. This coefficient is also comparable with most of the regression results in table 4.2. Also sales is significant at the 5% level and positively related to the price cost margin. The small differences between the results based on fixed effects and first-differences indicate that the outcomes are not severely affected by problems associated with the stationarity of the variables, model misspecification and measurement error.[36]

Specification 6 includes interaction terms between a dummy variable that indicates the type of sector (manufacturing or services) and the innovation input and output variable in a fixed effects model. This specification is used to investigate whether the effect of innovation output and input on competition are different for firms in the manufacturing and services sector.[37] In the services sector, the price cost margins of firms increase on average with 0.134 percentage point with each percentage point increase of the innovation output variable. This effect is lower in manufacturing sectors (0.059 percentage

[36] However, one drawback of specification 5 is that the number of observations is reduced due to the utilization of the First-Differences estimator. Another drawback is that innovation input and output are not instrumented. Instrumenting based on first-differences, a dummy that indicates marketing activities and the amount of R&D related personnel leads to insignificant results due to problems with weak instruments.

[37] The dummy variable that indicates the type of sector and the interaction terms are significant at the 1% level. The F-statistic is 3.85.

Table 4.4: Estimation results: robustness checks

Specification	(5) *FD*	(6) *Fixed effects*	(7) *2-step GMM*
O_{it}	0.038***	0.134***	0.067***
	(0.013)	(0.035)	(0.012)
I_{it}	-0.151**	-0.063	-0.141**
	(0.065)	(0.071)	(0.065)
$MANUFAC_{it}$		-0.185***	
		(0.064)	
$O_{it} * MANUFAC_{it}$		-0.074**	
		(0.037)	
$I_{it} * MANUFAC_{it}$		-0.143*	
		(0.082)	
$COOPCOMP_{it}$	0.010*	0.006	0.007
	(0.006)	(0.005)	(0.005)
$COOPOTHER_{it}$	-0.004	-0.005	-0.006
	(0.005)	(0.004)	(0.004)
$PRODONLY_{it}$	-0.001	-0.001	-0.002
	(0.004)	(0.004)	(0.004)
S_{it}	0.0002**	0.0001***	0.0001***
	(0.0001)	(0.00003)	(0.00003)
$INNOFREQ_{it}$	-0.004	-0.004	-0.005
	(0.004)	(0.004)	(0.004)
Observations	1,843	4,057	4,070
Hansen J statistic			0.262
			(0.877)

The dependent variable is the competition variable (price cost margin). Standard errors (between parentheses) are robust against heteroskedasticity and serial correlation. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01. The chi-squared statistic and p-value (between brackets) are reported for the Hansen test. Only instruments for the innovation input variable are used in specification 7.

point) and may be the result of a higher risk of imitation by competitors in manufacturing (Sirilli and Evangelista, 1998). Hence, while innovation output negatively affects competition in both the manufacturing sector and the services sector, competition is reduced more in services sectors. However, innovation input plays a more important role for competition in the manufacturing sector than in the services sector. Price cost margins decrease on average only by approximately 0.063 percentage point in the services sector if innovation input increases by one percentage point, while they decrease by 0.205 percentage point in the manufacturing sector. Thus, accumulated innovation experience has a stronger positive effect on competition in the manufacturing sector than in the services sector. An explanation for this result may be that innovating is more costly in manufacturing and therefore leads to a stronger effect of accumulated expenditures on price cost margins.

Specification 7 is similar to specification 3, but is based on a different measure of accumulated innovation experience. The innovation expenditure intensities are accumulated instead of accumulating the total innovation expenditures and normalizing by sales. Only the innovation input variable is instrumented based on R&D researchers, assistants and other R&D related personnel. The innovation output variable is significantly positive at the 1% level and implies that increases of one percentage point in the relative sales from new products or services lead to an increase of 0.067 percentage point in the price cost margin. The alternative measure for accumulated innovation experience is negative and significant at the 5% level. The magnitude of the coefficient (-0.141%) is comparable with most of the results in Table 4.2.

In sum, these robustness checks are consistent with the estimation results shown in table 4.2. Higher sales from new products relative to the total sales lead to less competition. More innovation input increases the probability to innovate and affect competition positively. The innovation output in the manufacturing sector has relatively less impact on competition compared to innovation output in the services sector. However, more accumulated innovation experience results in more competition in the manufacturing sector than in the services sector.

4.5 Conclusion

This chapter contributes to the literature on competition and innovation by examining the effect of innovation output and input on competition. Previous studies do not examine these effects in one framework and therefore cannot identify the relative importance of innovation output or input for the degree of competition. Furthermore, these studies may be subject to omitted variable bias due to the relationship between innovation output and input. A model of tacit collusion is used in chapter to show that firms become more dominant after they innovate and that an increase of the probability that an innovation occurs hinders collusion. Consequently, an innovation output reduces the degree of competition, while an innovation input that increases the probability to innovate leads to more competition. Two-yearly data from the Dutch Community Innovation Survey and Production Statistics over the period 1994-2004 are used. The results in this study show that firms with a higher proportion of sales from new products have higher price cost margins, while firms with more cumulative innovation expenditures (normalized by sales) have lower price cost margins. Hence, innovation output leads to less competition and innovation input results in more competition. Although the estimated coefficients of the innovation output and input variable are somewhat larger in absolute value when utilizing instruments, the results seem to be robust across different specifications. Furthermore, the findings in this chapter suggest that innovation output is more important for competition in the services sector relative to the manufacturing sector. However, innovation input plays a more important role in the manufacturing sector compared to the services sector. The results also provide evidence that innovation input is less important for competition than innovation output.

The findings in this chapter are also of interest for competition and innovation policies. The negative effect of innovation output on competition is important for competition policies due to feedback effects. According to Aghion et al. (2005b), the effect of competition on innovation is characterized by an inverted-U relationship. For example, if the initial degree of competition is low and authorities endeavor to increase the level of competition to enhance innovation, the negative effect of innovation output on competition has to be taken into account since this negative feedback effect on competition mitigates the initial competition stimulus and thereby lowers innovation. The results in this study are also important for patent policies and R&D subsidies since both policy instruments can be used to affect the degree of competition. Strict patent policies result in more sales from new products and lead to higher price cost margins of firms. While firms benefit from these innovations by gaining a competitive edge and capturing the rents from these innovations, competition in a market may be impaired due to strict patent policies. However, R&D subsidies enable firms that are eligible to receive these funds to conduct additional research which in turn increases their probability to innovate. According to the results, increases in the probability to innovate encourage firms to behave more competitively by hindering tacit collusion. Hence, policies aimed at increasing the chances that innovations occur in a sector may mitigate the anti-competitive effects of patent policies. The findings also suggest that policies that affect innovation output are relatively more important for the degree of competition compared to policies aimed at innovation input.

Chapter 5

Producing Innovations: Determinants of innovativity and efficiency

5.1 Introduction

The importance of technical change as a driving force of economic growth and prosperity has been widely recognized in the literature. E.g., (Aghion and Howitt, 1998, p. 151) state that "[T]he chances of achieving sustainable growth depend critically on maintaining a steady flow of technological innovations". There is a long tradition in the empirical research that has identified research and development (R&D) as a key input in generating this flow of technological innovations. A seminal contribution to this literature was made by Griliches (1980), who estimated a model for the US manufacturing industry, including R&D expenditures, to seek an explanation for the observed productivity slowdowns in 1965-73 and 1973-78. Indeed, Griliches (1980) finds that the productivity slowdown in the period 1965-73 was largely due to the collapse of R&D investment. These results inspired more studies confirming that R&D drives innovation at the the firm level (Griliches, 1986; Jaffe, 1986; Griliches, 1998), the industry level (Griliches and Lichtenberg, 1984; Nadiri, 1980) and across countries (Griliches and Mairesse, 1983, 1991; Mansfield, 1988).

From its early inception, researchers in this field have looked at the innovation process as others have looked at the production of goods and services and have estimated a knowledge production function (KPF). Using the analogy to the production function, the relation between R&D efforts and innovative outputs has been investigated with increasing econometric rigor and increasing levels of detail and quality in the data (Coe and Helpman, 1995; Park, 1995; Engelbrecht, 1997; Lichtenberg and van Pottelsberghe de la Potterie, 1998; Keller, 2002; Guellec and Van Pottelsberghe de la Potterie, 2004; Griffith et al., 2004). Looking at the process of knowledge production, Mairesse and Mohnen (2002) define *innovativeness* as the unexplained ability to turn innovation inputs into innovation output (analogous to (total factor) productivity in the production function). Inovativeness encompasses omitted factors such as technological, organizational, cultural, environmental factors or inefficiency (Mairesse and Mohnen, 2002). Following the traditional growth accounting logic, the changes in innovation output can then be ascribed to changes in the innovation inputs and a residual that picks up changes in innovativeness.

From this literature we know that not every R&D dollar or hour spent produces the same new knowledge and although innovation is strongly positively correlated with growth performance at all levels of aggregation, the link between R&D expenditures and economic growth cannot be understood without explaining these differences (Ulku, 2004; Acs and Audretsch, 1991). For example, the European 'innovation paradox' is based on the observation that, even after correcting for differences in R&D investments, countries in the European Union lag behind their US competitors in creating economic value from these investments (Figel, 2006). And less spectacular, but possibly more important from an innovation management perspective, there are large differences in innovation output among firms in an industry (Thompson, 2001; Cohen and Klepper, 1996; Cohen and Levin, 1989) even after controlling for R&D expenditures. In other words, there seems to be substantial heterogeneity in innovativeness. But much like the Solow-residual, this unexplained heterogeneity is a measure of our ignorance more than anything else.

The major methodological contribution of this chapter lies in taking the production function analogy one step further to explain this variation. Innovativeness, like productivity, can conceptually be split between (in)efficiency and technology (e.g. Weil (2008)). In productivity analysis it is quite common to separate (in)efficiency from technological change empirically using Stochastic Frontier Analysis (SFA) (e.g. Aigner et al. (1977); Battese and Corra (1977); Meeusen and van den Broeck (1977)). In the empirical literature on the KPF, however, researchers to date still assume, usually implicitly, that all innovation takes place at the frontier and no waste of R&D

inputs occurs.[1] To the best of our knowledge there is only a handful of papers out there (e.g. Wang (2007); Wang and Huang (2007); Fu and Yang (2009); Gantumur and Stephan (2010)) that has applied stochastic frontier analysis to the estimation of a KPF. The first three estimated the KPF at the country level and found that indeed inefficiencies pick up a substantial share of the cross country variation. Gantumur and Stephan (2010) is the only (unpublished) paper we have found to apply this method to micro-data. Their results show that the variance in inefficiency is about twice the variance of the remaining unexplained error in the German context.[2] We therefore feel that (in)efficiency should not be assumed away and that efficiency differences may go a long way in explaining differences in innovative performance. Moreover, not accounting for (in)efficiency (changes) and its determinants potentially biases the estimates of the parameters in the innovation function (Greene, 2005).[3]

A second contribution lies in our results. We estimate the knowledge production frontier using the Community Innovation Survey, an unbalanced firm level panel data set collected in the Netherlands between 1994 and 2004. This database allows us to span the entire innovation process from initial resources committed (R&D labor and the accumulated knowledge stock) to the final resulting sales volume of new products.[4] Our analysis allows us to ask and answer several important questions. The first and most basic ques-

[1]This is implicit in the assumption that observations are randomly scattered around the "true" knowledge production frontier that is estimated. The implication is that all (in)efficiency differences and changes end up in the error term. As Thompson (2001) puts it, there is only little evidence on differences in firm ability to innovate because these can only be observed indirectly. We fill that gap by presenting an established method to estimate innovative ability.

[2]Several others have employed the closely related Data Envelopment Analysis (DEA) to estimate the innovation frontier (see Zhang et al. (2003)). Coelli et al. (2005) provides a discussion of the differences between DEA and SFA. Like these DEA-papers, the Gantumur and Stephan (2010) paper focuses on the benchmarking of firm level R&D performance, overlooking the implications for the literature to which we intend to make our contribution.

[3]With the advent of endogenous growth theory (Romer, 1990; Aghion and Howitt, 1992; Grossman and Helpman, 1991) precisely estimating unbiased parameters of the KPF has gained importance from a theoretical point of view. The scale effects in the first generation endogenous growth models (Jones, 1995) depend on the value of the parameters in the innovation function and Ha and Howitt (2007) and Madsen (2008) estimated knowledge production functions to distinguish between second generation Schumpeterian (Aghion and Howitt, 1998; Dinopoulos and Thompson, 1998; Peretto, 1998; Young, 1998; Howitt, 1999) and semi-endogenous growth models (Jones, 1995; Kortum, 1997; Segerstrom, 1998).

[4]We excluded process innovations because typically one firms process innovation is another (upstream) firm's product innovation.

tion is: "Are innovation processes efficient and how important is inefficiency in the innovation process? We find that indeed (in)efficiency accounts for a significant part (depending on the specification between 50 and 92%) of the unexplained between firm and over time variation in innovation output. Furthermore, changes in efficiency explains on average 62% of innovativess as defined by Mairesse and Mohnen (2002).

Our data then allow us to explore the following research questions: "what is the impact of cooperative innovation activities on the efficiency and productivity of innovation inputs?" and: "what is the impact of government support on the productivity of innovation inputs and efficiency in the innovation process?" We find no efficiency differential between those that do and do not cooperate with their competitors but do find that firms receiving government support are more efficient than those that do not. This result can have many reasons. It may imply that only efficient innovators have the time and inclination to apply for support or that the government agencies selecting projects are doing a good job in finding the most efficient applicants (selection effect) or that subsidies force firms to be more efficient ex post (learning effect). We do not explore this issue further here. Based on the average level of knowledge stock and research labor, the innovativity (marginal productivity) of the knowledge stock and research labor do not differ between firms with and without cooperative innovation activities or between firms with and without government funding. By combining our data with information from the Dutch Production Statistics we can also relate (in)efficiencies in the innovation process to the firm's price-cost margin. If we interpret this margin as a measure of competition we find that more competitive firms are more innovative in terms of generating new product sales from innovations. We also find that larger firms are typically less efficient, suggesting that hierarchy and bureaucracy are bad for innovative efficiency.

The contribution of this chapter is to rigorously employ the tools developed for productivity analysis to the emerging field of 'innovativeness analysis'. Our study is one of the first that introduces stochastic frontier analysis in this literature. Gantumur and Stephan (2010) estimate a distance function and focus on different research questions. They focus on the acquisition of external technology, while our study focuses on cooperation with competitors/ other institutions and government funding. In addition we feel that our empirical results add new and important insights for theoretical and empirical researchers and for a more policy and management oriented readership. The remainder of the chapter is structured as follows. Section 5.2 describes the methodology used and data collected to estimate the innovation function. In Section 5.3, we describe our results. Section 5.4 concludes.

5.2 Methodology and Data

5.2.1 Methodology

We specify an innovation function and use stochastic frontier analysis to identify efficiency as a key component in innovativeness. SFA was first introduced by Aigner et al. (1977), Battese and Corra (1977), and Meeusen and van den Broeck (1977). It has been developed and applied in productivity analysis at the micro and macro-level.[5] Figure 5.1 is an example of an innovation frontier with a single innovation output and input. The innovation output function is concave due to diminishing returns to the innovation input.[6] In this example, firm (a), (b) and (c) are subject to the same innovation frontier and operate at different innovative output levels. Firms (a) and (c) operate on the innovation frontier without inefficiencies in their innovation process and firm (c) has less innovation output as it employs less innovation input. Firm (b) produces less innovations than (a) despite the fact that they use the same amount of innovation input. This discrepancy is attributed to the presence of inefficiencies in the innovation process of firm (b).

In actually estimating the knowledge production frontier we follow endogenous growth theory and assume that all firms use an accumulated knowledge stock.[7] In addition we want to control for firm size and the intensity of product market competition. The reason to control for size is that large and small firms report differently on their R&D inputs (e.g., Kleinknecht et al. (1991)) and this systematic measurement error might affect our estimations. The control for competition intensity is necessary because our proxy for innovation output, innovative product sales, may be systematically higher for firms in less competitive markets by the fact that market power allows them to charge higher prices for the same innovation output.

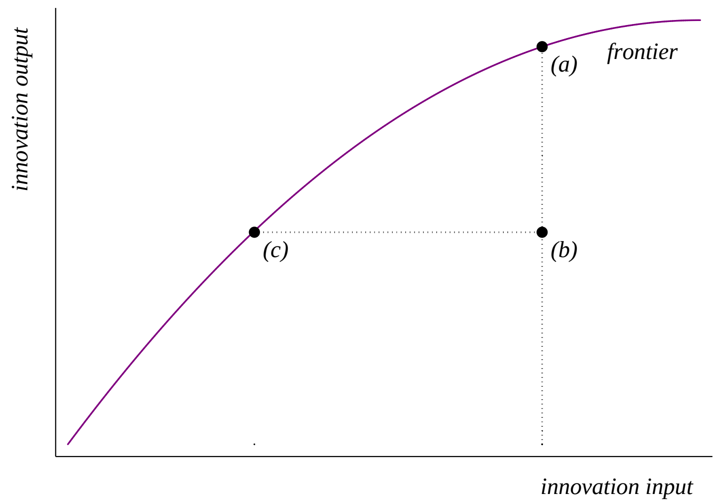
To answer our research questions we include dummies for cooperation

[5] Kumbhakar and Lovell (2000) provide an elaborate discussion of the development and application of SFA to efficiency measurement.

[6] We do not impose diminishing returns in the empirical model below. The figure merely serves to illustrate the methodology.

[7] See e.g. Jones (1995).

Figure 5.1: Innovation frontier



with competitors and other institutions, for funding from government agencies and for time (to check if indeed technical change in the knowledge production function is slow or absent). To also allow for flexibility in the functional form we estimate the frontier as a translog innovation function:

$$\begin{aligned}\ln Y_{it} = {} & \beta_K \ln K_{it} + \beta_L \ln L_{it} + \tfrac{1}{2}\beta_{KK} \ln K_{it}^2 + \tfrac{1}{2}\beta_{LL} \ln L_{it}^2 \\ & + \beta_{KL} \ln K_{it} \ln L_{it} + \tau_t D_t + \beta_z z_{it} + \beta_{Kz} \ln K_{it} z_{it} + \beta_{Lz} \ln L_{it} z_{it} \\ & + \tfrac{1}{2}\beta_{KKz} \ln K_{it}^2 z_{it} + \tfrac{1}{2}\beta_{LLz} \ln L_{it}^2 z_{it} + \beta_{KLz} \ln K_{it} \ln L_{it} z_{it} \\ & + \beta_C C_{it} + \beta_{FS} FS_{it} + \beta_j + \beta_i + v_{it} - u_{it},\end{aligned} \tag{5.1}$$

where Y_{it} is our innovation output variable, K_{it} represents the knowledge stock, L_{it} denotes the flow of R&D effort (research labor), D_t are time dummies that capture "technical change" in the innovation process, z_{it} is a vector of dummy variables to indicate whether firms cooperate or not and whether they receive funding or not, C_{it} is our measure of competition intensity, FS_{it} represents firm size, β_j are technology class dummies, β_i are firm specific random effects over time, $u_{it} \geq 0$ is assumed to be i.i.d., with a half-normal distribution truncated at zero, —$N(0, \sigma_u^2)$—, and independent from disturbance term v_{it} that follows a white noise process. By including time dummies in the specification, we allow for shifts in the innovation frontier in a more flexible manner than using a time trend (Baltagi and Griffin, 1988).[8] Dummies are included to control for the industry technology classes based on the OECD industry classification (See Raymond et al. (2009) appendix A).

After obtaining the estimated parameters of the innovation frontier we can compute the (in)efficiency for each firm. Their efficiency scores are measured as the ratio of actual over the maximum attainable innovation output if firms would innovate on the frontier, where $0 \leq \exp\{-u_{it}\} \leq 1$, and $\exp\{-u_{it}\} = 1$ implies full efficiency. We relate technical inefficiency u_{it} in the stochastic frontier model to our dummy variables and controls to investigate how (in)efficiency responds to these variables. Technical inefficiency u_{it} in the stochastic frontier model is specified as:

$$u_{it} = z_{it}\gamma_z + \gamma_C C_{it} + \gamma_{FS} FS_{it} + w_{it}, \tag{5.2}$$

The noise term w_{it} is defined by the truncation of the normal distribution with zero mean and variance σ_w^2. A one step model is used, where the specified stochastic frontier model in equation (5.1) and the endogenous inefficiency term in equation (5.2) are estimated in a single step by maximum likelihood.[9] The estimated coefficients, γ_z and $\gamma_{C,FS}$, now relate cooperation, funding, competition and firm size to the efficiency of the innovative process, i.e. the distance between firm (a) and (b) in Figure 5.1 is related to our variables of interest. Cooperation, government funding, competition and firm size can thus affect the position and shape of the frontier and the distance to the frontier.

In the basic specifications, we estimate the innovation function by specifying a Cobb-Douglas production function, a translog production function with and without explaining the inefficiency component in the innovation production process. The basic specifications are used to examine whether a Cobb-Douglas or a translog production function is more appropriate and to investigate the importance of inefficiency in the innovation production process. Mairesse and Mohnen (2002) emphasize the importance of what remains to be explained in the production of innovations (innovativeness or TFP). Based on the basic specifications, we also decompose productivity

[8] In contrast to studies that estimate a normal production function, when estimating the KPF there is no reason to assume a constant time trend. There is no reason to expect that "technological change" makes knowledge production more or less innovative over time in the same way that new technologies improve productivity in the production process.

[9] In two-step estimations, stochastic frontier models are estimated first and the relationship between inefficiency and covariates in the second step. Wang (2002) show that two-step estimations produce biased estimates.

change (innovativeness or TFP) by identifying (the share of) pure technical change, a scale component (knowledge stock and labor) and efficiency change to analyze the importance of inefficiency in the innovation production process. Finally, extended specifications are used by including interaction terms between the innovation inputs and dummy variables that represent cooperative innovation activities and funding from the government. The extended specifications are used to analyze whether cooperation with competitors, other institutions and funding from the government affect the innovativity (marginal productivity) of the knowledge stock or labor in the innovation production process.

5.2.2 Data

To estimate our innovation frontiers we use firm-level data on innovation from the Community Innovation Survey (CIS) in The Netherlands.[10] For the purpose of this study, the CIS data are merged with financial information from the Production Survey (PS). The CIS data contain information on the R&D and other innovation activities of the firm, such as innovation expenditures, innovation activities conducted with other institutions, the effects of the innovation output (e.g. quality improvement, product differentiation etc.) and sources of the knowledge used to produce innovations. The PS data provide information on output, employment, value added, profit and other financial information. Both the CIS and PS data are collected by the *Centraal Bureau voor de Statistiek*. The sample from the CIS is based on five survey waves, namely CIS 2 (1994-1996), CIS 2.5 (1996-1998), CIS 3 (1998-2000), CIS 3.5 (2000-2002) and CIS 4 (2002-2004). In The Netherlands, each innovation survey is conducted every two years. The CIS and PS data are a combination of census data and a stratified random sample. The census data contain all firms with 50 employees or more and the stratified sample is based on firms with less than 50 employees. The stratum variables are the economic activity and the number of employees, where the economic activity of a firm is based on the Dutch standard industrial classification (SBI). Firms that are included in one survey only are excluded from the sample.[11] The population of interest are firms with at least 10 employees and positive sales.

In the CIS questionnaire, firms are asked first to provide general information on their economic activity, sales, number of employees etc. The

[10] Brouwer et al. (2008) and Raymond et al. (2009) also use Dutch Community Innovation Survey.

[11] We assume that attrition of the panel data occurs exogenously.

second part of the questionnaire contains questions about the innovation activities of firms, such as their R&D activities, the percentage sales from new product/services, other innovation input expenditures, partnerships in innovation activities etc. Firms are asked to provide information on the second part of the CIS questionnaire if they affirm one of the three questions regarding: 1) whether firms developed new or strongly improved products 2) whether firms used new or strongly improved production processes 3) whether the firm has ongoing or abandoned innovation activities. Firms are classified as innovators if they affirm one of these three questions.

5.2.3 Innovation output

We use the sales from new or improved products as our measure of innovation output, Y_{it}.[12] The main advantage of this innovation output measure is that it captures innovations directly by measuring the introduction and the success of the newly developed products or services. Conventional innovation output measures such as patents or citation-weighted patents cannot capture the output of all innovations since many innovations are not patented and patented ideas are not always commercialized.[13] A drawback of the sales from new products as an innovation output measure is that firms may provide only rough estimates of their sales due to innovative products. This may induce measurement error in the regression analysis.[14] Another drawback of this measure is that the sales from new products may be influenced by the life cycle of a product.[15] In our analysis, firms with more than 50% of their sales from new products are therefore excluded from the sample.[16] Firms with only process innovations are also excluded since these

[12]The analysis in this chapter is restricted to products new to the firm instead of using an innovation output measure based on products new to the market. Brouwer et al. (2008) argue that a measure based on products new to the market may suffer from problems related to the interpretation of firms regarding their scope of the relevant market. This may lead to overestimation of innovation efforts by firms that are more focused at home markets.

[13]See Kamien and Schwartz (1982) and Geroski (1990) for a discussion on the limitations of patents as an innovation output measure.

[14]Measurement error in the dependent variable does not affect the consistency of the parameter estimates if the component that represents the deviation from the true value of the dependent variable is not correlated with the (composite) error term or explanatory variables.

[15]New product sales follow a logistic curve as the product diffuses in the market. This implies that (small) firms that do R&D and introduce a new product may experience the large increases in their new product sales with quite a lag.

[16]This cut-off point is also used by Raymond et al. (2009).

firms have no sales from innovations new to the firm. Hence, our sample consists of firms with only product innovations or both product and process innovations.

5.2.4 Innovation inputs

We follow Hall and Jones (1999) and construct the stock of total innovation expenditures by using a perpetual inventory method to proxy for K_{it}, the accumulated knowledge stock.[17] The stock of total innovation expenditures represents the knowledge capital of a firm. While many papers do not account for learning effects in patent races or the innovation process, it is assumed in this chapter that the knowledge stock or accumulated innovation experience is a primary input in the innovation process. Doraszelski (2003) shows that firms have incentives to reduce R&D expenditures if their knowledge stock increases. R&D efforts in the past affect the probability to win the R&D race positively.

To construct our stock, we use total innovation expenditures because using R&D expenditures will understate the knowledge base of small firms (and consequently bias the output elasticity of the other inputs up). In addition to internal and external R&D expenditures, total innovation expenditures also include the purchase of rights and licenses to use external technology, and the purchase of advanced machinery and computer hardware devoted to the implementation of product and process innovations. [18]

We define the knowledge stock as:

$$K_{it} = (1 - \delta)K_{it-1} + I_{it}, \tag{5.3}$$

where K_t is the knowledge stock and I_t represents the total innovation expenditures during period t. Furthermore, the depreciation rate is assumed to be 15% and the pre-sample growth rate of innovation expenditures 5%.[19] The knowledge stock at the beginning of the first period is defined by the following equation:

[17] Hall and Jones (1999) use R&D expenditures instead of total innovation expenditures.

[18] Total innovation expenditures is only based on these components in our study. The total innovation expenditures in the pre-CIS4 data also includes additional components. However, these components are not included as continuous measures in CIS4 and are excluded to ascertain the consistency of this measure across CIS waves.

[19] These values for the depreciation rate and growth rate are often used to construct the knowledge stock (Hall and Mairesse (1995)). The results are robust when a 10% and 20% depreciation rate for the knowledge stock are used.

$$
\begin{aligned}
K_{i1} &= I_{i0} + (1-\delta)I_{i-1} + (1-\delta)^2 I_{i-2} + ... \\
&= \sum_{s=0}^{\infty}(1-\delta)^s I_{i-s} = I_{i0}\sum_{s=0}^{\infty}\left(\frac{1-\delta}{1+g}\right)^s = \frac{I_{i0}}{g+\delta}
\end{aligned}
\tag{5.4}
$$

In addition to the stock of knowledge, we include the amount of researchers in R&D activities in full-time equivalents as a flow input, L_{it}, in the innovation process.

5.2.5 Determinants of innovativity and efficiency

In this study, we aim to examine whether firm size, competition, cooperation on innovation activities and innovation subsidies affect the (efficiency of the) innovation process. Firm size is measured by the total number of employees in the firm. And two cooperation dummy variables indicate whether firms cooperate with competitors (COOPCOMP) and other institutions (COOPOTHER) in their innovation activities. We also include a dummy variable to examine the effect of R&D funding from the government (FUNDING) on the innovativity and efficiency in the innovation process. The value 1 indicates firms with funding from the local government, national government or European Union. The reference group for this dummy variable consists of firms without R&D funding from the government. Competition intensity cannot be observed directly, but proxies have been suggested in the literature. As in Aghion et al. (2005b), we use the price cost margin (viz., the Lerner index, or markup) as our measure of competition. The price cost margin is calculated by dividing the total sales minus the cost of sales, labor expenses and energy costs by total sales:

$$C_{it} = (\frac{S_{it} - TVC_{it}}{S_{it}}), \tag{5.5}$$

where C_{it} is the competition variable, S_{it} total sales and TVC_{it} represents the total variable costs. We use a firm-level measure of competition, since industries are relatively broadly defined in the data set and the intensity of competition can differ between firms, even within narrowly defined industries.[20] Hence, we assume that all changes in competition are reflected in the price cost margins of firms.[21] An important advantage of the price cost

[20]The classification is based on the 3-digit SBI.

[21]An important assumption in this thesis is that less competition is positively related with higher price cost margins, while more competition leads to lower price cost margins.

margin over conventional measures of competition such as the Herfindahl-Hirschman Index (HHI) and concentration ratios is that the price cost margin does not require a precise definition of the relevant geographical or product boundaries. To eliminate outliers in our measure we eliminated observations that fall outside the range -1 and 1.

Table 5.1 provides the descriptive statistics of innovation output, innovation inputs, cooperative innovation activities, funding from the government, the price cost margin and the number of employees. The means of the sales from innovations, the knowledge stock and research labor in fte are €7,481,56; €3,466,236 and 2.442 fte, respectively. Most of the firms are not cooperating on innovation activities according to our data. Only 11.8% of the firms cooperate with competitors on innovation activities, 37.8% cooperate with other institutions, while the remaining 60.6% of the firms are not cooperating on their innovation activities. More than half of the firms in the sample receive funding from a local/ regional authority, the central government or the European Union, namely 63.8%. On average, firms earn a price cost margin of 24.8% and have 187 employees on their payroll.

Table 5.1: Descriptive statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
Sales from innovations in €1,000	7,481.567	14,319.87	1.521	218,986.6
Knowledge stock in €1,000	3,466.236	4,885.73	18.61	28,876.23
Research labor in fte	2.442	3.965	0.029	40
Cooperation with competitors	0.118	0.323	0	1
Cooperation with other institutions	0.378	0.485	0	1
Funding from the government	0.638	0.481	0	1
Price cost margin	0.248	0.1119	-0.816	0.704
Number of employees (own payroll)	187.04	350.678	0	10,857

The descriptive statistics are based on the sample in Column v in Table 5.2 (1,366 observations).

5.3 The results

5.3.1 Basic specifications

Table 5.2 presents the results based on the basic specifications. Column i shows a basic Cobb-Douglas specification estimated with SFA and allowing for inefficiency. To examine the presence of inefficiency in the innovation process, a likelihood-ratio test is performed assuming the null hypothesis of no technical inefficiency ($H_0 : \sigma_u = 0$). The null hypothesis is rejected at the 1% level and indicates the presence of inefficiency in the innovation process. In this basic specification, we find that (in)efficiency accounts for approximately 22% of the variation in the residual (the ratio of variation in (in)efficiency σ_u over total variation $\sigma_u + \sigma_v$) .[22] Both innovation input variables are individually significant at the 1% level. The output elasticities of the knowledge capital stock and research labor in this specification are 0.43 and 0.16, respectively. The innovation function has decreasing returns to scale since $\beta_K + \beta_L < 1$. Our basic specification, however, only hints at inefficiencies in the production of innovations, but does not explain possible causes of inefficiencies.

Table 5.2: The innovation function: basic specifications

Panel A: Basic parameters of the Cobb-Douglas and Translog Specification				
Specification	(i) Cobb-Douglas	(ii) Translog	(iii) Translog	(iv) Translog
$\ln K_{it}$	0.431***	-0.033	0.176	-0.065
	(0.031)	(0.325)	(0.363)	(0.267)
$\ln L_{it}^{R\&D}$	0.161***	-0.213	-0.957***	0.042
	(0.033)	(0.264)	(0.296)	(0.221)
$\frac{1}{2}\ln^2 K_{it}$		0.059	-0.004	0.111*
		(0.044)	(0.049)	(0.058)
$\frac{1}{2}\ln^2 L_{it}^{R\&D}$		-0.031	-0.174***	-0.010
		(0.042)	(0.048)	(0.035)
$\ln K_{it} \ln L_{it}^{R\&D}$		0.054	0.178***	0.005
		(0.035)	(0.039)	(0.049)
$COOPCOMP_{it}$				-0.020
				(0.127)
$COOPOTHER_{it}$				0.154*
				(0.086)
$FUNDING_{it}$				0.037
				(0.081)
C_{it}				-1.361***
				(0.342)
FS_{it}				0.005***
				(0.001)
Panel B: Determinants of inefficiency				
$COOPCOMP_{it}$			-0.528***	0.155
			(0.176)	(0.736)
$COOPOTHER_{it}$			-0.184**	-0.238

Continued on next page ...

[22] We also estimated the basic specifications with industry dummies instead of dummies for the technology classes. The inclusion of these dummies has a negligible effect on the parameter estimates.

Table 5.2 (Continued from previous page)

			(0.091)	(0.529)
$FUNDING_{it}$			-0.072	-0.727
			(0.082)	(0.485)
C_{it}			0.319	-3.405*
			(0.420)	(1.969)
FS_{it}			-0.003***	0.006***
			(0.0001)	(0.001)
$\sigma_u/(\sigma_u + \sigma_v)$	0.220	0.989	0.499	0.918
Observations	1,367	1,367	1,366	1,366
Technology class	yes	yes	yes	yes
Technical change	no	no	no	yes

The dependent variable is sales from innovations. Standard errors (between parentheses) are robust against heteroskedasticity. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01.

Column ii provides the results based on a translog innovation function. The translog specification is a generalization of the Cobb-Douglas function and an F-test shows that it is preferred over the Cobb-Douglas specification in column i.[23] Column ii shows that all variables are individually insignificant in the translog specification. However, all terms are jointly significant at the 1% level. The output elasticities with respect to the knowledge stock and labor are 0.41 and 0.18, respectively.[24] On average, these output elasticities do not differ that much from output elasticities based on the Cobb-Douglas specification in column i. Unlike column i, however, the specification in column ii attributes 99% of the variance to inefficiencies in the innovation process.[25] Both specifications do not explain possible sources of inefficiency.[26]

Column iii shows the results based on a translog innovation function, where inefficiency is related to cooperation with other competitors, cooperation with other institutions, funding from the government, competition and firm size. However, we assume that the innovation frontier is the same for all firms and do not allow for innovativity differences between firms with re-

[23]The F-test is performed on the quadratic terms and the interaction terms. The Cobb-Douglas is not rejected if the coefficients of these terms are equal to zero. The translog function is used in further specifications since the Cobb-Douglas specification is rejected at the 1% significance level.

[24]These output elasticities are evaluated at the average (natural logarithm of) knowledge capital and research labor.

[25]We get this very high ratio because relative to a Cobb-Douglas, the translog specification is much more flexible and allows for a substantial reduction in the variance of v.

[26]The null hypothesis of no technical efficiency is again rejected at the 1%.

spect to the knowledge stock and labor. Only the research labor, its squared term and the interaction term are individually and jointly significant at the 1% level. The knowledge stock, its squared term and the interaction term are jointly significant at the 1% level. The output elasticities with respect to the knowledge stock and labor are 0.16 and 0.33, respectively.[27] These output elasticities differ from the output elasticities in Column i and ii. The marginal innovativity of the knowledge stock is lower, while the marginal innovativity of the research labor is higher compared to previous specifications. Furthermore, in contrast to the results in Column i and ii, the output elasticity of research labor is higher than the output elasticity of the knowledge stock in column iii. Compared to the previous specification, the variance attributed to (in)efficiency is lower in column iii since it drops from 99% to 50%.

The results in column iii (Panel B) show the determinants of inefficiency. Cooperation with competitors (COOPCOMP) is significant at the 1% level and is negatively related to inefficiency. This means that firms that cooperate with competitors are on average more efficient. Cooperation with other institutions (COOPOTHER) is significant at the 5% level and negatively related to inefficiency. Firms that cooperate with other institutions are on average more efficient than firms without cooperative innovation activities. Funding from the government (FUNDING) is not related with inefficiency in the innovation production process. Cooperation with competitors is more important for inefficiency compared to cooperation with other institutions. The findings in column iii show no evidence of a relationship between price cost margins and inefficiency in the production of innovations. Firm size (FS) is significantly negatively related to inefficiency at the 1% level. This means that larger firms are producing innovations more efficient (on average). This finding is consistent with the outcomes of Gantumur and Stephan (2010). Large firms may be more efficient if they use more specialized inputs in production. However, this specification assumes that all firms are subject to the same innovation frontier and does not allow for differences in the innovativity of the knowledge stock and labor between firms with and without cooperation on innovation activities or firms with and without government funding. Furthermore, the specifications in column i-iii do not include firm size and competition in the innovation function directly.

Column iv presents the results based on including cooperation with competitors, other institutions, competition and firm size directly in the in-

[27] These output elasticities are evaluated at the average (natural logarithm of) knowledge capital and research labor.

novation function and as explanatory variables for the inefficiency term.[28] Again, we observe a rise in the ratio of variance in (in)efficiency over total variance (to 92%), because introducing more controls in the specification of the frontier implies that the overall fit is improved with the same variance in (in)efficiency. Although the knowledge capital stock, research labor, quadratic terms and interaction term are individually insignificant, they are jointly significant at the 1% level. The output elasticities with respect to knowledge capital and labor are 0.24 and 0.08, respectively.[29] The output elasticities are lower than the output elasticities from the Cobb-Douglas and translog specification in Column i and ii, where competition, cooperation with competitors, cooperation with other institutions, funding from governments and firm size are not included in the innovation function and inefficiency is assumed to be exogenous. While the marginal innovativity of the knowledge stock is higher than in the specification in Column iii, the marginal innovativity of labor is smaller. Cooperation with competitors and funding from the government are not significantly related to innovation output. Only cooperation with other institutions is positively related to the innovation output, but only at the 10% significance level. The price cost margin is significant at the 1% level and negatively related to innovation output. This means that more competition is positively related to the sales from innovations. A possible explanation is that competitive pressures may induce firms to introduce more successful innovations.[30] Another explanation is that new innovations cannibalize the sales from existing products and therefore lower the price cost margins. Firms size is significant at the 1% level and positively related to innovation output. An explanation for this result is that larger firms have developed larger distribution and marketing networks to sell their new products.

Cooperation with competitors, cooperation with other institutions and funding from governments are not significantly related to inefficiency in the innovation process. These results are not consistent with the findings in column iii since cooperation with competitors and other institutions are no longer significantly related to inefficiency after the inclusion of these vari-

[28] We also estimated this specification with interaction terms between the innovation inputs and technology classes to examine whether the marginal innovativity of the knowledge stock and research labor differ between technology classes (industry groups). The interaction terms are not jointly significant. This suggests that the marginal innovativity of the knowledge stock and research labor do not differ across technology classes.

[29] These output elasticities are evaluated at the average (natural logarithm of) knowledge capital and research labor.

[30] A positive relationship between price cost margins and the sales from innovations is expected if less competition allows firms to reap the rents from their innovations.

ables in the innovation function. Competition is significantly related to inefficiency at the 10% level, where firms with higher price cost margins have lower inefficiency. In contrast to the results in column iii, Firm size is significantly positively related to inefficiency at the 1% level. Larger firms may suffer from coordination problems and therefore experience on average more inefficient innovation processes. Jorde and Teece (1990) argue that in a simultaneous model of innovation, tight linkages and feedback mechanisms which must operate quickly and efficiently must also exist within firms.[31] These findings with respect to firm size are also not in line with the findings of Gantumur and Stephan (2010). A possible explanation for the discrepancy between the findings is that firm size is not directly included in the innovation function in column ii and in the specification used by Gantumur and Stephan (2010). A drawback of the specification in Column iv is that the interaction between the innovativity of knowledge capital, labor and cooperation with institutions and funding from governments cannot be examined. These interactions are necessary to examine whether the marginal innovativity of the knowledge stock and labor differ between these groups and are examined in section 5.3.3. In the next section, the importance of inefficiency in the innovation production process is examined by decomposing productivity change.

5.3.2 Decomposing productivity change

To examine the importance of inefficiency for innovativeness (as defined by Mairesse and Mohnen (2002)) and the production of innovations, we decompose productivity change into technical change, technical efficiency change and the contribution of returns to scale. The change in innovatiness or productivity change is defined as $\dot{INN} = \dot{Y} - \dot{X}$.[32] We use the basic specification in Column ii and the decomposition by Kumbhakar and Lovell (2000):

$$\dot{INN} = T\Delta + (\theta - 1) \sum_{n} \left(\frac{\theta_n}{\theta} \right) \dot{x} + TE\Delta, \tag{5.6}$$

where $\dot{INN}$ represents the change in innovativeness (productivity change), $T\Delta$ is the (pure) technical change, θ is the sum of the output elasticities over n inputs, θ_n is the output elasticity of an innovation input, $\dot{x}$ is the change in

[31] For example, innovation requires efficient coordination and feedback between the research and marketing department.

[32] The dot over the variables indicates the rate of change.

an innovation input and $TE\Delta$ is the change in technical efficiency. Table 5.3 provides an overview of the average change in innovativeness, pure technical change, returns to scale components and technical efficiency. Furthermore, the average shares of the decomposition components are shown.

Table 5.3: Decomposing the change in innovativeness

Variable	Average change	Share
$T\Delta$	-0.018	13.099%
$(\theta-1)\left(\frac{\theta_K}{\theta}\right)\dot{K}$	-0.021	15.465%
$(\theta-1)\left(\frac{\theta_L}{\theta}\right)\dot{L}$	-0.012	8.746%
$TE\Delta$	-0.085	62.691%
$I\dot{N}N$	-0.136	100%

The decomposition of the productivity change is based on Column ii in Table 5.2. The share of each decomposition component in explaining productivity changes is based on the average change in the decomposition components.

The results in Table 5.3 show that innovativeness (productivity change) contributes negatively to the growth in innovation output, namely -13.6% on average. Pure technical change, the change of the scale components with respect to the knowledge stock and research labor, and the change in technical inefficiency are -1.8%, -2.1%, -1.2% and -0.085%, respectively. The results in Table 5.3 also show that inefficiency accounts on average for 62.69% in explaining the change in innovativeness (productivity change). Hence, the negative effect of the change in innovativeness on the growth of innovation output is mostly attributable to inefficient innovation production processes. The remaining 37.31% of the negative effect of the change in innovativeness of innovation output is due to (negative) changes in pure technical change and returns to scale with respect to the knowledge stock and research labor. Therefore, it is important to examine the determinants of inefficiency and productivity change.

5.3.3 Extended specifications

In column v we show the results based on the interaction between the innovation inputs and the dummy variables that indicate whether firms cooperate with competitors or other institutions and whether they receive funding from the government. As in specification iv, the variance in inefficiency explains 92% of the total variance of the composite residual. F-tests are used to examine whether the innovativity (marginal productivity) of the knowledge stock and research labor differs between cooperating firms and firms that do not cooperate and firms with and without funding. The cooperation with competitors dummy and interaction terms with the knowledge stock and research labor are not jointly significant. Also the dummy variable that indicates cooperation with other institutions and interaction terms are not jointly significant. However, an F-test based on funding from the government and interaction terms indicates that there are significant differences in the innovativity of the knowledge stock and labor between firms with funding from the government and firms without funding.[33] The output elasticities of the knowledge stock for firms without funding and with funding are significant at the 1% level and 0.21 and 0.27, respectively.[34] The output elasticities of the research labor for firms without funding and with funding are significant at the 5% level and 0.18 and 0.09, respectively.[35] While the point estimates of the output elasticity with respect to the knowledge stock is somewhat higher for firms with funding from the government compared to firms without funding, the innovativity of research labor is lower for firms with funding. We also performed t-tests to examine whether the output elasticities differ significantly between firms with and without funding. There are no significant differences in innovativity between firms with and without funding based on the t-tests.[36]

Table 5.4: The innovation function: extended specifications

Panel A: Basic parameters of the Cobb-Douglas and Translog Specification		
Specification	(v) Translog	(vi) Translog
$\ln K_{it}$	0.849*	-0.578
	(0.459)	(0.415)
$\ln L_{it}^{R\&D}$	-0.595*	0.129
	(0.359)	(0.371)
$\frac{1}{2}\ln^2 K_{it}$	0.042	-0.089
	(0.036)	(0.063)
$\frac{1}{2}\ln^2 L_{it}^{R\&D}$	-0.018	-0.011
	(0.054)	(0.068)
$\ln K_{it} \ln L_{it}^{R\&D}$	0.106**	0.004

Continued on next page ...

[33]The dummy variable and interaction terms are significant at the 5% level.

[34]The output elasticities are evaluated at the average natural logarithm of the knowledge stock and research labor.

[35]The output elasticities are evaluated at the average natural logarithm of the knowledge stock and research labor.

[36]The t-values with respect to the knowledge stock and research labor are 1.07 and 1.47, respectively.

Table 5.4 (Continued from previous page)

	(0.029)	(0.047)
Panel B: Cooperation with competitors & the innovativity of K and L		
$COOPCOMP_{it}$	-7.975**	-1.980
	(3.865)	(4.562)
$\ln K_{it} * COOPCOMP_{it}$	2.149**	0.496
	(1.007)	(1.227)
$\ln L_{it}^{R\&D} * COOPCOMP_{it}$	-1.117	0.947
	(0.796)	(0.952)
$\frac{1}{2}\ln^2 K_{it} * COOPCOMP_{it}$	-0.283**	-0.070
	(0.130)	(0.163)
$\frac{1}{2}\ln^2 L_{it}^{R\&D} * COOPCOMP_{it}$	-0.143	0.297**
	(0.125)	(0.146)
$\ln K_{it} \ln L_{it}^{R\&D} * COOPCOMP_{it}$	0.138	-0.153
	(0.102)	(0.125)
Panel C: Cooperation with other institutions & the innovativity of K and L		
$COOPOTHER_{it}$	3.023	-0.635
	(2.522)	(3.057)
$\ln K_{it} * COOPOTHER_{it}$	-0.789	0.496
	(0.685)	(1.227)
$\ln L_{it}^{R\&D} * COOPOTHER_{it}$	0.353	0.947
	(0.577)	(0.952)
$\frac{1}{2}\ln^2 K_{it} * COOPOTHER_{it}$	0.103	-0.070
	(0.092)	(0.163)
$\frac{1}{2}\ln^2 L_{it}^{R\&D} * COOPOTHER_{it}$	-0.032	0.297**
	(0.084)	(0.146)
$\ln K_{it} \ln L_{it}^{R\&D} * COOPOTHER_{it}$	-0.055	-0.145
	(0.076)	(0.125)
Panel D: Funding from the government & the innovativity of K and L		
$FUNDING_{it}$	4.742**	-0.852*
	(1.965)	(0.464)
$\ln K_{it} * FUNDING_{it}$	-1.387**	0.098
	(0.538)	(0.066)
$\ln L_{it}^{R\&D} * FUNDING_{it}$	1.023**	0.933**
	(0.444)	0.372
$\frac{1}{2}\ln^2 K_{it} * FUNDING_{it}$	0.200***	-0.002
	(0.073)	(0.003)
$\frac{1}{2}\ln^2 L_{it}^{R\&D} * FUNDING_{it}$	0.076	0.173**
	(0.071)	(0.081)
$\ln K_{it} \ln L_{it}^{R\&D} * FUNDING_{it}$	-0.153***	-0.142***
	(0.059)	(0.049)
Panel E: Other controls that affect innovation output		
C_{it}	-1.332***	-1.512***
	(0.340)	(0.416)

Continued on next page ...

Table 5.4 (Continued from previous page)

FS_{it}	0.005***	0.006***
	(0.001)	(0.001)
Panel F: Determinants of inefficiency		
$COOPCOMP_{it}$	-0.133	-1.043*
	(0.723)	(0.597)
$COOPOTHER_{it}$	-0.237	-0.283
	(0.522)	(0.375)
$FUNDING_{it}$	-0.682	-0.912***
	(0.472)	(0.341)
C_{it}	-3.289*	-2.344*
	(1.876)	(1.251)
FS_{it}	0.005***	0.006***
	(0.001)	(0.001)
$\sigma_u/(\sigma_u+\sigma_v)$	0.920	0.825
Observations	1,366	926
Technology class	yes	yes
Technical change	yes	yes

The dependent variable is sales from innovations. Standard errors (between parentheses) are robust against heteroskedasticity. The results in Column vi are based on a lagged effect of co-operation with competitors, other institutions and funding. Asterisks indicate significance at the following levels: * – 0.10, ** – 0.05, and *** – 0.01.

When we examine the determinants of inefficiency we conclude that only competition and firm size are on average related to inefficiency. The relationship between competition and inefficiency in the innovation process is quite weak since it is only significant at the 10% level. Higher price cost margins (less competition) are related to lower inefficiency. Consistent with the findings in column iv, firm size is significantly positively related with inefficiency at the 1% level. Coordination problems in the production of innovations by large firms may lead to less efficient innovation production processes. Cooperation with competitors or other institutions and funding from governments are on average not related to inefficiency.[37]

Column vi provides an overview of the estimation results based on a

[37]However, when we calculate efficiency scores and examine the differences in the distributions between groups, we find that the distributions of the efficiency scores differ between firms with and without funding from governments. A Kolmogorov-Smirnov test is used and the distributions differ significantly from each other at the 1% significance level. Based on a kernel density, we find that at higher efficiency scores, there are more firms with funding from governments than without funding. Furthermore, there are more firms without funding at lower levels of efficiency scores. Hence, firms with funding from the government are more likely to produce innovations efficiently compared to firms without funding from the government. We do not find differences in the distributions of efficiency scores between firms with and without cooperative innovation agreements.

lagged effect of cooperative innovation agreements and government subsidies.[38] The variance of the inefficiency term accounts for most part of the total variance of the residual, namely 83%. With respect to cooperation on innovation activities, the F-tests indicate that the innovativity of the knowledge stock and research labor differs between firms that cooperate and firms that do not cooperate.[39] Furthermore, the tests show that there are differences between firms with funding from the government and firms without funding. However, evaluated at the average natural logarithm of the knowledge stock and research labor, we do not find that the point estimates of the output elasticities differ significantly between firms with and without cooperation on innovation and with and without government funding.[40] The price cost margin is significant at the 1% level and negatively related to the sales from innovations. This finding is consistent with the results in the previous specifications. Firm size is significant at the 1% level and larger firms have more sales from new innovations.

We also examined the lagged effect of cooperation with competitors, cooperation with other institutions and funding from the government on inefficiency in the innovation production process. Cooperation with competitors in the previous period is negatively related with inefficiency at the 10% level. Thus we find weak evidence that cooperation with competitors in the past leads to improved efficiency in the future. There is no significant relationship between cooperation with other institutions and inefficiency. Funding from the government is significantly related to inefficiency at the 1% significance level. Firms that received funding from the government in the previous period are on average more efficient in the current period. Competition is significant at the 10% level and as in earlier specifications, less competition is positively related to efficiency. Moreover, larger firms are significantly less efficient than smaller firms.

[38]The lagged effect corresponds to a delayed effect of two years since bi-annual data are used.

[39]Cooperation with competitors and the interaction terms are significant at the 5% level. Cooperation with other institutions and the interaction terms are significant at the 1% level.

[40]This conclusion is based on t-tests to examine whether each output elasticity differs from the output elasticity of firms without cooperation and without funding. The absolute t-values range between 0.05 and 1.15. Nevertheless, differences in output elasticities may also arise due to different levels of the knowledge stock and labor. The change in the knowledge stock differs significantly at the 5% level between firms that received funding only in the previous period compared to firms that received no funding. The average change in the knowledge stock is 462,000 euros for firms with funding and 76,000 euros for firms without funding. However, we do not find significant differences in the output elasticities between these types of firms evaluated at their average knowledge stock.

5.4 Conclusion

This chapter contributes to the innovation literature by examining several sources of innovativity and efficiency in the innovation process. We use Stochastic Frontier Analysis and estimate innovation frontiers for Dutch firms over the period 1994-2004. Dutch Community Innovation Survey data is used to examine whether cooperation with competitors, cooperation with other institutions, funding from the government, competition and firm size affect innovativity and efficiency in the production of innovations. We find that inefficiency is present in the innovation process of Dutch firms and that percentage changes in efficiency explains on average 63% of changes productivity.

We find that cooperation with competitors and funding from the government are on average significantly negatively related to inefficiency in the basic specification where the innovation frontier is assumed to be homogenous for all firms. There is no relationship on average between cooperation with competitors, funding from the government and ineffiency once we allow these factors to affect innovation output directly and the innovativity of the knowledge stock and research labor. However, the distribution between efficiency scores of firms with and without government funding seem do differ significantly, where relatively more firms with high efficiency scores have government funding and relatively more firms with low efficiency scores have no government funding. Funding from the government affects the innovativity of the knowledge stock and labor significantly, but the magnitude of the output elasticities do not seem to differ between groups. For example, firms with and without funding have the same innovativity of the knowledge stock and labor evaluated at the average knowledge stock and research labor. We also examine lagged effects of cooperation with competitors, other institutions and government funding on the innovativity of innovation inputs and inefficiency in the innovation production process. Also in the specification with lagged effects, we do not find differences in innovativity between firms with and without cooperative innovation activities and between firms with and without government funding. However, while there seems to be only weak evidence of a positive effect of cooperation with competitors in the past on future efficiency, firms that receive funding from the government are on average more efficient in future periods. In all specifications, we find that competition is significantly positively related to a higher level of innovation output. However, there is only weak evidence that competition affects inefficiency in the innovation process. The results also suggest that while larger firms have higher levels of innovation output, their innovation process seems to be subject to more inefficiencies.

Further research should aim at examining the channels through which cooperation with competitors, funding from the government and other factors affect the innovativity and efficiency in the innovation process. For example, funding from the government may improve the innovativity of the knowledge stock and research labor, but it may also be the case that firms with a high innovativeness are better able to attract funds from the government. Furthermore, it is not clear whether more efficient firms receive these funds or whether firms with funds from the government become more efficient. If government funds lead to more efficient innovation production processes, this policy instrument can be used to aid firms in their quest to introduce successful innovations. Future research should also aim at examining the channel through which firm size affects inefficiency in the innovation production process. For example, if large firms suffer from coordination problems between departments during the production of innovations, firm policies that enhance communication and coordination between the departments may lead to improvements in innovation production processes.

Chapter 6

Conclusions

The analyses in this thesis are focused on the relationship between competition and innovation. In particular, a new measure of innovation is introduced to investigate the effect of competition on innovation in financial services, the effect of innovation on competition is examined and the innovation production process is explored. This chapter contains a recapitulation of the main findings presented in this thesis. Section 6.1 provides a summary of the findings per chapter and the policy implications. Finally, section 6.2 discusses important limitations of the analyses in this thesis and provides some directions for future research.

6.1 Summary of the findings and policy implications

A new measure of innovation is introduced in chapter 2 to analyze the relationship between competition and innovation in the U.S. banking industry. The extant literature on financial innovation and determinants is scarce since conventional innovation measures that are often used in studies on manufacturing (such as patents and R&D expenditures) are not obvious in financial services. The theoretical model of Aghion et al. (2005b) is applied on the banking industry to explain the competition-innovation relationship in U.S. banking. The theoretical model incorporates both positive and negative effects of competition on innovation that are often found in the literature to explain the existence of an inverted-U relationship between competition and innovation. The empirical analysis in chapter 2 is based on individual bank Call Report data for the period 1984-2004 and consists of two parts. First, a new measure of overall innovation is introduced by estimating and enveloping annual minimum cost frontiers to create a global frontier. The distance to this global frontier captures the technology gap of banks which in turn is influenced by their innovation behavior. Second, the empirical analysis focuses on estimating the competition-innovation relationship in U.S. banking and exploring how certain market developments, such as consolidation and interstate banking deregulation, affect competition and innovation. Consistent with the findings of Aghion et al. (2005b) and others, the outcomes in this thesis suggest that there is an inverted-u relationship between competition and innovation in U.S. banking. Increased competitive pressures enhance innovation at lower levels of competition, while more competition is detrimental for innovation at higher levels of competition. The outcomes in chapter 2 also show that price cost margins increased on average between 1984-2004, a period of substantial consolidation. Furthermore, the interstate banking deregulation, that allowed banking systems to integrate across states, led on average to less competition. Since the price cost margin of most banks is above their optimal price cost margin that enhances innovation, increases in the price cost margins due to the deregulation led to less innovation.

In chapter 3, both the theoretical and empirical analysis in chapter 2 and Aghion et al. (2005b) are extended to examine the competition-innovation nexus in a consolidating market more closely. The analysis centers around the question whether banks are 'too big to innovate'. While Aghion et al. (2005b) assume that unit costs are independent from the quantity produced, the model in chapter 3 allows for U-shaped average cost curves as is often found in the banking industry. Thus firms may experience scale economies or diseconomies of scale. In the theoretical model in chapter 3, general conditions are derived under which the innovation behavior of firms with scale diseconomies becomes more or less responsive to changes in competition. As in chapter 2, technology gaps between banks are estimated and utilized as an innovation measure. Based on individual bank data over the period 1984-2004, the findings in chapter 3 show that banks experienced more diseconomies of scale over time while the average bank size and price cost margin increased. More important, the results not only show the existence of an inverted-U relationship between competition and innovation, but also that the innovation by large banks is significantly more sensitive to changes in competition.

The analyses in chapter 2 and 3 provide more insights into the competition-innovation nexus and related market developments in U.S. banking, such as the interstate banking deregulation and ongoing consolidation. Banks operate mostly at price cost margins that are higher than the optimal price cost margin that enhances innovation. These results imply that developments that lead to the diminution of competition could negatively affect the innovation incentives of banks in the future. Consistent with the findings in other studies, this paper shows that the interstate banking deregulation affected price cost margins positively. One explanation provided in the literature is that the deregulation triggered the reallocation of resources from less profitable banks towards banks with higher price cost margins. In light of the competition-innovation nexus in U.S. banking, this deregulation affected innovation negatively via its effect on competition. Furthermore, the findings in this thesis show that less competition lowers innovation by large banks more than proportionally. During the recent financial crisis, many governments intervened in the banking industry to stabilize the financial system via mergers of large failing banks and other financial institutions. Given the consolidation of large troubled banks, future changes in competition will predominantly affect the innovation behavior of these large institutions. However, the results in this thesis also imply that policies aimed at increasing the degree of competition in U.S. banking lead to more innovation by banks as they attempt to escape from their competitors. Especially large banks benefit from such competition policies since their innovation behavior is more responsive to changes in competition.

While chapter 2 and 3 focus on the effect of competition on innovation, chapter 4 provides more insights into the effect of innovation on competition by separating two dimensions from innovation, namely innovation input and output. The effects of these dimensions on competition are examined in one framework to identify the relative importance of these dimensions for competition. A model of tacit collusion is used and the probability to innovate is linked to accumulated innovation experience as an innovation input. Increases in accumulated innovation experience positively affect the probability to innovate and hinder tacit collusion. Firms that expect to innovate successfully and to surpass their competitors are less likely to collude tacitly as they fear less from retaliations against their competitive behavior. Hence, more innovation input affects competition positively. In addition, the model shows that once a firms innovates successfully, the innovator surpasses the competitors and increases its dominance in the market. Therefore, innovation output leads to higher price cost margins and less competitive behavior. Bi-annual Dutch Community Innovation Survey data over the period 1994-2004 are used to investigate the effect of innovation input and output on competition. The results in chapter 4 suggest that more successful innovation output leads to less competition, while increases in innovation efforts lead to more competition. The effect of innovation output on competition is larger in the service sector than in the manufacturing sector. An explanation for this result is that the risk of imitation is a more important hampering factor to innovation in manufacturing than in services. Innovation input has a larger effect on competition in the manufacturing sector than in services. Innovation input is more costly in manufacturing and may have a stronger effect on competition compared to the service sector. The empirical findings also show that innovation output is a relatively more important determinant of competition than innovation input.

In chapter 2 and 3, the effect of competition on innovation is examined and the policy implications are discussed. The findings in these chapters clearly show that competition policies can be used to enhance the innovation incentives of firms. At lower levels of competition, more competition stimulates firms to become more innovative in an attempt to escape from their competitors. However, the results in chapter 4 indicate that an initial competition stimulus may be mitigated if successful innovators become more dominant in the market. Hence, policymakers should take possible feedback effects into account when implementing competition policies to foster innovation. The results in chapter 4 are also of interest for policies related to patents and available innovation subsidies from governments. Patents provide firms with the opportunity to exploit the rents from their innovations by providing some protection against imitation. Strict patent policies lead to higher rents from innovation and a stronger effect of innovation output on competition by increasing the dominance of successful innovators. Hence, patent policies are an important tool to reward innovations and to affect the degree of competition. Moreover, the outcomes in chapter 4 show that conducive innovation environments with respect to innovation input may affect the degree of competition positively. For example, R&D subsidies provide firms the opportunity to conduct additional research that would otherwise not be conducted without the funding. Firms may behave more competitively if successful innovations become more likely due to the increased innovation efforts. Hence, policies with respect to innovation input may mitigate the negative effects of patent policies on the degree of competition. Nevertheless, the findings in chapter 4 also imply that policies aimed at innovation output are more important for the degree of competition than policies with respect to innovation input.

Chapter 4 shows that innovation input and output both have implications for competition and that the effects of these dimensions on competition should be examined in one framework due to the interrelatedness between innovation input and output. Chapter 5 contains a more detailed analysis of the relationship between innovation output and input. Specifically, the innovation production process is examined by estimating innovation functions and by investigating the determinants of productivity and efficiency. Sales from new innovations are used as an innovation output measure and the stock of knowledge and research labor are used as innovation inputs. The results show that the innovation production function exhibits decreasing returns to scale and that inefficiencies are present during the production of innovations. Furthermore, the decomposition analysis shows that changes in efficiency explain around 63% of changes in productivity (innovativeness). Furthermore, the variation in inefficiency explains between 50% and 99% of the total variation in the composite residual (depending on the specification). Hence, efficiency plays an important role in the innovation production process. With regard to the innovation function, the results show that there are no significant differences in the productivity of the innovation inputs between firms with and without cooperative innovation activities. Furthermore, there are no significant differences in the productivity of the innovation inputs between firms with and without government funding. More competition and larger firms are positively related to more sales from innovations. With regard to the determinants of efficiency in the innovation production process, the results show that there is only a weak contemporaneous relationship between competition and efficiency. Furthermore, the outcomes show that larger firms are less efficient. While there is no contemporaneous relationship between cooperation on innovation activities, funding from the government and inefficiency, the findings indicate that there is a positive lagged effect of funding from the government on efficiency.[1]

Inefficiency is present in the innovation production process and constitutes an important part of this process. Hence, improvements in the efficiency of this process fosters the production of innovations. Since large firms seem to be less efficient than their smaller counterparts, large firms should investigate the causes of this inefficiency to enhance their sales from innovations. An important finding for government policy in chapter 5 concerns the relationship between the lag of government funding and inefficiency. If innovation subsidies from the government in previous periods affect efficiency positively in later periods, these policy instruments seem to be effective by fostering innovation output directly and by reducing the slack in the innovation process.

[1] There is only a weak relationship between the lag of cooperation with competitors and inefficiency.

6.2 Limitations and suggestions for future research

The analyses in this thesis are subject to several important limitations, but provide a basis for some interesting research directions. In chapter 2, technology gaps are introduced as a new innovation measure to examine the relationship between competition and innovation in U.S. banking. An important drawback of this approach is that it captures innovation behavior indirectly. Annual minimum cost frontiers are estimated and enveloped to create a global frontier that represents the best potentially available technology. While conventional measure such as patents and R&D expenditures may not capture all innovation behavior of firms, technology gaps may capture other factors than innovations only. Drawbacks of conventional measures is that not all innovations are patented and require R&D expenditures. Hence, these conventional measures provide underestimations of the innovation behavior of firms. However, Total Factor Productivity (TFP) captures more than innovation behavior due to influences of inefficiency and scale economies on TFP. In the estimation procedure to obtain technology gap ratios, the influence of inefficiency and scale economies are corrected for. These corrections provide an advantage in the use of technology gap ratios over measures based on TFP. However, some other factors than innovation may still affect the technology gaps. For example, unequal access by firms to the labor force may influence the technology set of firms and may therefore affect technology gaps. An interesting direction for future research is the relationship between other measures of innovation and technology gap ratios. For example, while banks do not have R&D budgets, they do have Information Technology (IT) budgets at their disposal. IT expenditures capture partly the total innovation incentives of banks. Hence, while it does not capture all innovations by banks, it measures innovation incentives directly.

Another drawback of the technology gap ratios based on cost frontiers is that it captures mostly process innovations. It captures only product innovations to the extent that these product innovations lead to cost reductions. Several important products that were introduced during the last decades are mortgage-backed securities, interest rate swaps, currency swaps, credit default swaps and internet banking. The recent financial crisis raised concerns with respect to certain product innovations that entailed high risks. The complexity of certain derivatives products makes it more difficult to assess risks associated with these investments. New innovations may foster the performance of banks and financial development, but higher risks due to these innovations may also jeopardize financial stability. Future research may aim at exploring the relationships between competition, the introduction of new products, performance, risks and financial stability.

In chapter 3, the effect of scale economies on the competition-innovation relationship is examined. General conditions are derived in this thesis to explain whether the innovation behavior of large firms are more or less responsive to competition. An interesting research direction is the improvement of the theoretical model by modeling the interaction between scale economies and competition more explicitly in the model. Modeling the interaction between scale economies and competition may aid in the quest to derive clear testable hypotheses from the model. For example, the theoretical model could be extended by allowing for the entry of other firms and by relating scale economies to the degree of competition. Scale economies may impose entry barriers and diminish the degree of competition by making entry more difficult. A limitation of the empirical analysis in chapter 3 is that the innovation measure and the scale economies measure from Stochastic Frontier Analysis (SFA) are obtained from similar estimation procedures. The result is a complex form of endogeneity in the empirical model that makes the interpretation of the results more difficult. The obtained results may be partly due to correlations by construction of the variables. If other innovation measures that are not based on SFA become available in financial services, similar analyses can be conducted without this specific endogeneity problem. These new analyses may ease the quantification of the economic impact of scale economies on the competition-innovation relationship.

Chapter 4 investigates the effect of innovation input and output on competition. The theoretical section in this chapter contains a basic model of tacit collusion where innovation input is related to the probability to innovate and the sustainability of tacit collusion. While the theoretical model shows tacit collusion becomes more difficult as the occurrence of an innovation becomes more likely, future research should aim at exploring more channels through which innovation may affect tacit collusion. For example, process innovations provide successful innovators cost advantages. These innovators may become more dominant due to their cost advantages, but cost asymmetries between firms may also hinder tacit collusion. Successful process innovators are less likely to collude since the retaliatory capabilities of non-innovators against competitive behavior are reduced due to their high costs. Innovations may also affect tacit collusion through the link between entry barriers, product differentiation, product quality improvements and tacit collusion. More theoretical and empirical research on these channels provides more information on the effects of the innovation on competition.

Chapter 5 examines the innovation production function and the determinants of the productivity of innovation inputs and efficiency. An interesting research direction in light of the findings in chapter 5 is an analysis that focuses on causal effects of the cooperative innovation activities, government subsidies, competition and firm size on efficiency in the innovation production process. Studies in the existing (scarce) literature on this topic are not able to capture the causal effects of the determinants of inefficiency. For example, capturing the causal effect of government funding on the productivity and inefficiency of innovation production processes is interesting for authorities to assess the successfulness of this policy instrument for the enhancement of innovation behavior of firms. Furthermore, it is interesting to examine the channels through which several determinants affect productivity and efficiency. For example, the results in this thesis suggest that larger firms are less efficient than their smaller counterparts. A possible explanation is that these large firms may suffer from coordination problems. Future research should focus on the channels through which firm size affects efficiency since it may provide interesting insights for recommendations with regard to firm policies. Related to the issue of causality and channels, an interesting research direction is to examine the possible existence of selection mechanisms with regard to productivity, inefficiency and determinants such as cooperation and government funding. For example, a positive relationship between the productivity of the innovation inputs, efficiency and cooperation with competitors may not be due to a positive effect of cooperation on productivity and efficiency. It is also possible that more productive and efficient innovators are more likely to engage in cooperative agreements. Finally, another interesting suggestion for future research is to investigate the dynamics with regard to the innovation production function. Over the years, more and detailed micro data on innovation has become available. For example, while this thesis only uses Community Innovation Survey data for five CIS waves over the years 1994-2004, additional CIS waves in the future allow for the examination of the dynamics of innovation. Innovation inputs may have a contemporaneous effect on innovation output, but another possibility is that delayed effects may exist. Also with regard to the determinants of efficiency and productivity. For example, cooperation with competitors and funding from the government may not have an instantaneous effect on productivity and efficiency, but may become effective after several years. Furthermore, it also enables an analysis on the persistency of inefficiency in the innovation production process. If inefficient innovation production processes are very persistent, it may be difficult for firms to improve their efficiency in the short-run.

References

Acemoglu, D. (2002). Directed technical change. *Review of Economic Studies 69*(4), 781–809.

Acs, Z. J. and D. B. Audretsch (1987a). Innovation in large and small firms. *Economics Letters 23*(1), 109–112.

Acs, Z. J. and D. B. Audretsch (1987b). Innovation, market structure, and firm size. *The Review of Economics and Statistics 69*(4), 567–574.

Acs, Z. J. and D. B. Audretsch (1991). R&D, firm size, and innovative activity. In Z. J. Acs and D. B. Audretsch (Eds.), *Innovation and Technological Change: An International Comparison*, University of Michigan Press, Ann Arbor, pp. 39–59.

Aghion, P., B. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005a). Competition and innovation: An inverted-U relationship. *NBER Working Paper 9269*.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005b). Competition and innovation: An inverted-U relationship. *Quarterly Journal of Economics 120*(2), 701–728.

Aghion, P. and R. Griffith (2005). *Competition and Growth: Reconciling Theory and Evidence*. Cambridge, MA: MIT Press.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies 68*(3), 467–492.

Aghion, P., C. Harris, and J. Vickers (1997). Competition and growth with step-by-step innovation: An example. *European Economic Review 41*(3-5), 771–782.

Aghion, P. and P. Howitt (1992). A model of growth through creative destruction. *Econometrica 60*(2), 323–351.

Aghion, P. and P. Howitt (1998). *Endogenous Growth Theory*. Massachusetts, USA: MIT Press.

Agrell, P. J., P. Bogetoft, and J. Tind (2002, July). Incentive plans for productive efficiency, innovation and learning. *International Journal of Production Economics 78*(1), 1–11.

Ahn, S. (2002). Competition, innovation and productivity growth. *OECD Economics Department Working Paper 317*.

Aigner, D. J., K. C. Lovell, and P. Schmidt (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics 6*(1), 21–37.

Akhavein, J., W. S. Frame, and L. J. White (2005). The diffusion of financial innovations: An examination of the adoption of small business credit scoring by large banking organizations. *The Journal of Business 78*(2), 577–596.

Albk, S. and L. Lambertini (1998). Collusion in differentiated duopolies revisited. *Economics Letters 59*(3), 305–308.

Allen, F. and D. Gale (1994). *Financial Innovation and Risk Sharing*. Cambridge, MA: MIT Press.

Altunbas, Y., E. P. M. Gardener, P. Molyneux, and B. Moore (2001). Efficiency in European banking. *European Economic Review 45*(10), 1931–1955.

Altunbas, Y., J. Goddard, and P. Molyneux (1999). Technical change in banking. *Economics Letters 64*(2), 215–221.

Baltagi, B. H. and J. M. Griffin (1988). A general index of technical change. *Journal of Political Economy 96*(1), 20–41.

Baptista, R. and P. Swann (1998). Do firms in clusters innovate more? *Research Policy 27*(5), 525–540.

Battese, G. E. and G. S. Corra (1977). Estimation of a production frontier model, with application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics 21*(3), 169–179.

Battese, G. E., D. S. P. Rao, and C. J. O'Donnell (2004). A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *Journal of Productivity Analysis 21*(1), 91–103.

Belderbos, R., M. Carree, and B. Lokshin (2004). Cooperative r&d and firm performance. *Research Policy 33*(10), 1477–1492.

Berger, A. N. (2003). The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, Credit and Banking 35*(2), 141–76.

Berger, A. N., A. Demirguc-Kunt, R. Levine, and J. Haubrich (2004). Bank concentration and competition: An evolution in the making. *Journal of Money, Credit and Banking 36*(3), 433–451.

Berger, A. N., R. S. Demsetz, and P. E. Strahan (1999, February). The consolidation of the financial services industry: Causes, consequences, and implications for the future. *Journal of Banking and Finance 23*(2-4), 135–194.

Berger, A. N. and R. DeYoung (2006). Technological progress and the geographic expansion of the banking industry. *Journal of Money, Credit and Banking 38*(6), 1483–1513.

Berger, A. N. and T. H. Hannan (1989). The price-concetration relationship in banking. *Review of Economics and Statistics 71*(2), 291–299.

Berger, A. N., A. K. Kashyap, and J. M. Scalise (1995). The transformation of the U.S. banking industry: What a long, strange trips it's been. *Brookings Papers on Economic Activity 26*(1995-2), 55–218.

Berger, A. N. and L. J. Mester (1997). Efficiency and productivity change in the U.S. commercial banking industry: A comparison of the 1980s and 1990s. *Federal Reserve Bank of Philadelphia Working Paper 97–5*.

Berger, A. N. and L. J. Mester (2003). Explaining the dramatic changes in performance of U.S. banks: Technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation 12*(1), 57–95.

Bernstein, D. (1996). Asset quality and scale economies in banking. *Journal of Economics and Business 48*(2), 157–166.

Bhattacharyya, S. and V. Nanda (2000). Client discretion, switching costs, and financial innovation. *Review of Financial Studies 13*(4), 1101–1127.

Bleaney, M. and K. Wakelin (2002). Efficiency, innovation and exports. *Oxford Bulletin of Economics and Statistics 64*(1), 3–15.

Boot, A. W. and A. V. Thakor (1997). Banking scope and financial innovation. *Review of Financial Studies 10*(4), 1099–1131.

Bos, J. W. B., J. W. Kolari, and R. van Lamoen (2009). Competition and innovation: Evidence from financial services. *TKI Working Paper 09-16, Tjalling Koopmans Institute, Utrecht School of Economics*.

Bos, J. W. B. and H. Schmiedel (2007). Is there a single frontier in a single European banking market? *Journal of Banking and Finance 31*(7), 2081–2102.

Brambor, T., W. Clark, and M. Golder (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis 14*(1), 63–82.

Brouwer, E., T. Poot, and K. Van Montfort (2008). The innovation threshold. *De Economist 156*, 45–71.

Carletti, E., P. Hartmann, and G. Spagnolo (2007). Bank mergers, competition, and liquidity. *Journal of Money, Credit and Banking 39*(5), 1067–1105.

Celent (2009). IT spending in financial services: A global perspective. News article 1009, Money Science, http://www.moneyscience.com/Technology_News/article1009.

Cetorelli, N. and P. E. Strahan (2006). Finance as a barrier to entry: Bank competition and industry structure in local U.S. markets. *The Journal of Finance 61*(1), 437–461.

Chang, M.-H. (1991). The effects of product differentiation on collusive pricing. *International Journal of Industrial Organization 9*(3), 453–469.

Chourchane, M., D. Nickerson, and R. J. Sullivan (2002). Financial innovation, strategic real options, and endogenous competition: Theory and an application to internet banking. *Federal Reverve Bank of Philadelphia, Conference on Innovation in Financial Services and Payments May.*

Coe, D. T. and E. Helpman (1995). International R&D spillovers. *European Economic Review 39(5)*, 859–887.

Coelli, T., D. P. Rao, and G. E. Battese (2005). *An Introduction to Efficiency Analysis* (2 ed.). New York: Springer.

Cohen, W. H. and R. Levin (1989). Empirical studies of innovation and market structure. In R. Schmalensee and R. Willig (Eds.), *The Handbook of Industrial Organization*, North-Holland, pp. 1060–1107.

Cohen, W. M. and S. Klepper (1996). A reprise of size and R&D. *The Economic Journal 106*(437), 925–951.

Cole, R. A., L. G. Goldberg, and L. J. White (2004). Cookie cutter vs. character: The micro structure of small business lending by large and small banks. *Journal of Financial and Quantitative Analysis 39*(2), 227–251.

Dearing, A. (2007). Enabling europe to innovate. *Science 315*(5810), 344–347.

Deneckere, R. (1983). Duopoly supergames with product differentiation. *Economics Letters 11*(1-2), 37–42.

DeYoung, R. and W. C. Hunter (2001). Deregulation, the internet, and the competitive viability of large banks and community banks. Working Paper Series WP-01-11, Federal Reserve Bank of Chicago.

Diewert, W. E. (1976, May). Exact and superlative index numbers. *Journal of Econometrics 4*(2), 115–145.

Dinopoulos, E. and P. Thompson (1998). Schumpeterian growth without scale effects. *Journal of Economic Growth 3*, 313–335.

Doraszelski, U. (2003). An R&D race with knowledge accumulation. *The RAND Journal of Economics 34*(1), 20–42.

Elyasiani, E. and S. M. Mehdian (1990). A nonparametric approach to measurement of efficiency and technological change: The case of large U.S. commercial banks. *Journal of Financial Services Research 4*(2), 157–168.

Engelbrecht, H.-J. (1997). International R&D spillovers, human capital and productivity in OECD economies: An empirical investigation. *European Economic Review 41*(8), 1479 –1488.

Farber, S. (1981). Buyer market structure and R&D effort: A simultaneous equations model. *The Review of Economics and Statistics 63*(3), 336–345.

Feuerstein, S. (2005). Collusion in industrial economicsa survey. *Journal of Industry, Competition and Trade 5*, 163–198.

Figel, J. (2006, November 21). The European response on the innovation paradox. Conference "Innovation paradox: the Flemish response? the EU response".

Frame, W. S. and L. J. White (2004). Empirical studies of financial innovation: Lots of talk, little action? *Journal of Economic Literature 42*(1), 116–144.

Fraser, D. R., J. W. Kolari, S. Pynnönen, and T. K. Tippens (2011). Market power, bank megamergers, and the welfare of bank borrowers. *Journal of Financial Research 33*, forthcoming.

Freixas, X. and J. Rochet (1997). *Microeconomics of Banking*. Cambridge: MIT Press.

Fu, X. and Q. G. Yang (2009). Exploring the cross-country gap in patenting: A stochastic frontier approach. *Research Policy 38*(7), 1203–1213.

Funke, M. and H. Strulik (2000). On endogenous growth with physical capital, human capital and product variety. *European Economic Review 44*(3), 491–515.

Gaba, V., P. Y., and U. G.R. (2002). Timing of entry in international market: An empirical study of U.S. fortune 500 firms in China. *Journal of International Business Studies 33*(1), 39–55.

Gallaher, M. and J. Petrusa (2006). Innovation in the U.S. service sector. *The Journal of Technology Transfer 31*(6), 611–628.

Gantumur, T. and A. Stephan (2010). Do external technology acquisitions matter for innovative efficiency and productivity? Technical Report 222.

Geroski, P. A. (1990). Innovation, technological opportunity, and market structure. *Oxford Economic Papers 42*(3), 586–602.

Geroski, P. A. and R. Pomroy (1990). Innovation and the evolution of market structure. *The Journal of Industrial Economics 38*(3), 299–314.

Gollop, F. M. and D. W. Jorgenson (1980). U.S. productivity growth by industry, 1947-73. In J. W. Kendrick and B. N. Vaccara (Eds.), *New Developments in Productivity Measurement and Analysis*, pp. 17–124. Chicago: University Chicago Press (for NBER).

Greene, W. H. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics 126*(2), 269–303.

Griffith, R., S. Redding, and J. V. Reenen (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *The Review of Economics and Statistics 86*(4), 883–895.

Griliches, Z. (1980). R&D and the productivity slowdown. *National Bureau of Economic Research Working Paper Series No. 434*.

Griliches, Z. (1986). Productivity, R&D, and the basic research at the firm level in the 1970's. *American Economic Review 76*(1), 141–54.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature 28*(4), 1661–1707.

Griliches, Z. (1998). Productivity and R&D at the firm level. In *R&D and Productivity: The Econometric Evidence*, pp. 100–133. NBER.

Griliches, Z. and F. Lichtenberg (1984). R&D and productivity growth at the industry level: Is there still a relationship? In Z. Griliches (Ed.), *R&D, Patents and Productivity*, Chicago. University of Chicago Press.

Griliches, Z. and J. Mairesse (1983). Comparing productivity growth: An exploration of french and U.S. industrial and firm data. *European Economic Review 21*(1-2), 89–119.

Griliches, Z. and J. Mairesse (1991). R&d and productivity growth: Comparing Japanese and U.S. manufacturing firms. Technical Report 1778.

Grossman, G. and E. Helpman (1991). *Innovation and Growth in the Global Economy.* Massachusetts, USA: MIT Press.

Guellec, D. and B. Van Pottelsberghe de la Potterie (2004). From R&D to productivity growth: Do the institutional settings and the source of funds of R&D matter? *Oxford Bulletin of Economics and Statistics 66*(3), 353–378.

Ha, J. and P. Howitt (2007). Accounting for trends in productivity and R&D: A Schumpeterian critique of semi-endogenous growth theory. *Journal of Money, Credit and Banking 39*(4), 733–774.

Hall, B. H. and J. Mairesse (1995). Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics 65*(1), 263–293.

Hall, R. E. and C. I. Jones (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics 114*(1), 83–116.

Hannan, T. H. and J. M. McDowell (1984). The determinants of technology adoption: The case of the banking firm. *Rand Journal of Economics 15*(3), 328–335.

Hashmi, A. (2007). Competition and innovation: The inverted-U relationship revisited. *Working paper, University of Toronto*.

Hauswald, R. and R. Marquez (2003). Information technology and financial services competition. *Review of Financial Studies 16*(3), 921–948.

Hayami, Y. and V. W. Ruttan (1970). Agricultural productivity differences among countries. *American Economic Review 60*(5), 895–911.

Howitt, P. (1999). Steady endogenous growth with population and R&D inputs growing. *The Journal of Political Economy 107*(4), 715–730.

Hughes, J. P. and L. J. Mester (1993). A quality and risk-adjusted cost function for banks: Evidence on the "too-big-to-fail" doctrine. *Journal of Productivity Analysis 4*(3), 293–315.

Humphrey, D. B. (1993). Cost and technical change: Effects from bank deregulation. *Journal of Productivity Analysis 4*(1-2), 9–34.

Hunter, W. C. and S. G. Timme (1991). Technological change in large U.S. commercial banks. *Journal of Business 64*(3), 339–362.

Ivaldi, M., B. Jullien, P. Rey, P. Seabright, and J. Tirole (2003). The economics of tacit collusion. *European Commission Final Report for DG Competition*.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review 76*(5), 984–1001.

Jones, C. I. (1995). R&D-based models of economic growth. *Journal of Political Economy 103*(4), 759–84.

Jones, K. D. and T. Critchfield (2005). Consolidation in the U.S. banking industry: Is the "long, strange trip" about to end? *FDIC Banking Review 17*(4), 31–61.

Jorde, T. M. and D. J. Teece (1990). Innovation and cooperation: Implications for competition and antitrust. *The Journal of Economic Perspectives 4*(3), 75–96.

Jovanovic, B. and G. M. MacDonald (1994). The life cycle of a competitive industry. *The Journal of Political Economy 102*(2), 322–347.

Kamien, M. I. and N. L. Schwartz (1982). *Market Structure and Innovation.* New York, NY, USA: Cambridge University Press.

Keller, W. (2002). Geographic localization and international technology diffusion. *American Economic Review 92*(1), 120–142.

King, R. G. and R. Levine (1993). Finance and growth: Schumpeter might be right. *Quarterly Journal of Economics 108*(3), 717–737.

Kleinknecht, A., T. P. Poot, and J. O. N. Reijnen (1991). Formal and informal R&D and firm size. survey results from the Netherlands. In Z. J. Acs and D. B. Audretsch (Eds.), *Innovation and Technological Change: An International Comparison*, Ann Arbor: University of Michigan Press, pp. 84–108.

Klepper, S. and K. L. Simons (2000). The making of an oligopoly: Firm survival and technological change in the evolution of the U.S. tire industry. *The Journal of Political Economy 108*(4), 728–760.

Kortum, S. S. (1997). Research, patenting, and technological change. *Econometrica 65*(6), 1389–1419.

Kroszner, R. S. and P. E. Strahan (1999). What drives deregulation? Economics and politics of the relaxation of bank branching restrictions. *Quarterly Journal of Economics 114*(4), 1437–1467.

Krugman, P. (1979). A model of innovation, technology transfer, and the world distribution of income. *Journal of Political Economy 87*(2), 253–266.

Kumbhakar, S. C. and C. A. K. Lovell (2000). *Stochastic Frontier Analysis*. Cambridge University Press.

Lau, L. J. and P. A. Yotopoulos (1989). The meta-production function approach to technological change in world agriculture. *Journal of Development Economics 31*(2), 241–269.

Levin, R. C., W. M. C. and D. C. Mowrey (1985). R&D appropriability, opportunity, and market structure: New evidence on some Schumpeterian hypotheses. *American Economic Review 75*(2), 20–24.

Levine, R. (1997). Financial development and economic growth: Views and agenda. *Journal of Economic Literature 35*(2), 688–726.

Levine, R. (2004). Finance and growth: Theory and evidence. Working Paper Series No. 10766, National Bureau of Economic Research.

Levine, R., N. Loayza, and T. Beck (2000). Financial intermediation and growth: Causality and causes. *Journal of Monetary Economics 46*, 31–77.

Levine, R. and S. Zervos (1998). Stock markets, banks, and economic growth. *American Economic Review 88*(3), 537–558.

Lichtenberg, F. R. and B. van Pottelsberghe de la Potterie (1998). International R&D spillovers: A comment. *European Economic Review 42*(8), 1483–1491.

Ludwig, E. A. (1997). Oral statement before the subcommittee on capital markets, securities and government sponsored enterprises at the U.S. house of representatives.

Lunn, J. (1986). An empirical analysis of process and product patenting: A simultaneous equation framework. *The Journal of Industrial Economics 34*(3), 319–330.

Madsen, J. (2008). Semi-endogenous versus Schumpeterian growth models: testing the knowledge production function using international data. *Journal of Economic Growth 13*, 1–26.

Mairesse, J. and P. Mohnen (2002). Accounting for innovation and measuring innovativeness: An illustrative framework and an application. *The American Economic Review 92*(2), 226–230.

Mansfield, E. (1988). Industrial R&D in Japan and the United States: A comparative study. *American Economic Review 78*(2), 223–28.

Mantel, B. and T. McHugh (2001). Competition and innovation in the consumer payments market? considering demand, supply, and public policy issues. *Federal Reserve Bank of Chicago Emerging Markets Working Paper EPS-2001-4*.

Meeusen, W. and J. van den Broeck (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review 18*(2), 435–444.

Melvin, M. and M. P. Taylor (2009). The crisis in the foreign exchange market. *Journal of International Money and Finance 28*(8), 1317–1330.

Miller, M. (1986). Financial innovation: The last twenty years and the next. *Journal of Financial and Quantitative Analysis 21*(4), 459–471.

Mishkin, F. S. and P. E. Strahan (1999). What will technology do to financial structure? In *NBER Working Papers*, Number 6892 in NBER Papers in Monetary Economics, pp. 41. National Bureau of Economic Research, Inc.

Mukhopadhyay, A. K. (1985). Technological progress and change in market concentration in the U.S., 1963-77. *Southern Economic Journal 52*(1), 141–149.

Mundlak, Y. and R. Hellinghausen (1982). The intercountry agricultural production function: Another view. *American Journal of Agricultural Economics 64*(4), 664–672.

Nadiri, M. I. (1980). Sectoral productivity slowdown. *American Economic Review 70*(2), 349–52.

O'Donnell, C., D. Rao, and G. Battese (2008). Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics 34*(2), 231–255.

Pagano, M. (1993). Financial markets and growth: An overview. *European Economic Review 37*(2-3), 613–622.

Park, K. and G. Pennacchi (2009). Harming depositors and helping borrowers: The disparate impact of bank consolidation. *The Review of Financial Studies 22*, 1–40.

Park, W. G. (1995). International R&D spillovers and OECD economic growth. *Economic Inquiry 33*(4), 571–591.

Peretto, P. F. (1998). Technological change and population growth. *Journal of Economic Growth 3*(4), 283–311.

Pilloff, S. J. (1999). Multimarket contact in banking. *Review of Industrial Organization 14*(2), 163–182.

Raymond, W., P. Mohnen, F. Palm, and S. Schim van der Loeff (2009). Innovative sales, R&D and total innovation expenditures: Panel evidence on their dynamics. Technical Report 028.

Raymond, W., P. Mohnen, F. Palm, and S. S. van der Loeff (2010). Persistence of innovation in Dutch manufacturing: Is it spurious? *Review of Economics and Statistics 92*(3), 495–504.

Rhoades, S. A. (2000). Bank mergers and banking structure in the United States, 1980-98. *Board of Governors of the Federal Reserve System Staff Studies 174*.

Romer, P. M. (1990). Endogenous technological change. *The Journal of Political Economy 98*(5), 71–102.

Ross, S. A. (1989). Institutional markets, financial marketing, and financial innovation. *Journal of Finance 44*(3), 541–556.

Rothschild, R. (1992). On the sustainability of collusion in differentiated duopolies. *Economics Letters 40*(1), 33–37.

Ruttan, V. W. (1997). Induced innovation, evolutionary theory and path dependence: Sources of technical change. *The Economic Journal 107*(444), 1520–1529.

Scherer, F. M. (1967). Market structure and the employment of scientists and engineers. *The American Economic Review 57*(3), 524–531.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy.* New York: Harper and Row.

Segerstrom, P. S. (1998). Endogenous growth without scale effects. *The American Economic Review 88*(5), 1290–1310.

Sirilli, G. and R. Evangelista (1998). Technological innovation in services and manufacturing: results from Italian surveys. *Research Policy 27*(9), 881–899.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics 39*(3), 312–320.

Stiroh, K. J. and P. E. Strahan (2003). Competitive dynamics of deregulation: Evidence from U.S. banking. *Journal of Money, Credit and Banking 35*(5), 801–828.

Subramanian, A. and S. Nilakanta (1996, December). Organizational innovativeness: Exploring the relationship between organizational determinants of innovation, types of innovations, and measures of organizational performance. *Omega 24*(6), 631–647.

Symeonidis, G. (1996). Innovation, firm size and market structure: Schumpeterian hypotheses and some new themes. *OECD Economics Department Working Paper 161.*

Tether, B. S. (2002). Who co-operates for innovation, and why: An empirical analysis. *Research Policy 31*(6), 947–967.

Thompson, P. (2001). The microeconomics of an R&D-based model of endogenous growth. *Journal of Economic Growth 6*(4), 263–283.

Timmer, M. P. and G. De Vries (2009). Structural change and growth accelerations in Asia and Latin America: A new sectoral data set. *Cliometrica 3*(2), 165–190.

Tinbergen, J. (1942). Zur theorie der langfristigen wirtschaftsentwicklung. *Weltwirtschafliches Archiv 55*(1), 511–549.

Tirole, J. (1988). *The Theory of Industrial Organization.* Cambridge, MA: The MIT Press.

Ulku, H. (2004). R&D, innovation, and economic growth: An empirical analysis. Technical Report 04/185.

Van Horne, J. C. (1985). Of financial innovations and excesses. *Journal of Finance 40*(3), 620–631.

Vives, X. (2001). Competition in the changing world of banking. *Oxford Review of Economic Policy 17*(4), 535–547.

Wang, E. C. (2002). Public infrastructure and economic growth: a new approach applied to east asian economies. *Journal of Policy Modeling 24*(5), 411–435.

Wang, E. C. (2007). R&D efficiency and economic performance: A cross-country analysis using the stochastic frontier approach. *Journal of Policy Modeling 29*(2), 345–360.

Wang, E. C. and W. Huang (2007). Relative efficiency of R&D activities: A cross-country study accounting for environmental factors in the DEA approach. *Research Policy 36*(2), 260–273.

Weil, D. N. (2008). *Economic Growth: International Edition*. New York, NY: Pearson Education.

Wheelock, D. C. and P. W. Wilson (1999). Technical progress, inefficiency, and productivity change in U.S. banking, 1984-1993. *Journal of Money, Credit and Banking 31*(2), 212–234.

Wilhelm, W. J. J. (2001). The internet and financial market structure. *Oxford Review of Economic Policy 17*(2), 235–247.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics 126*(1), 25–51.

Wu, C. and K. J. Wei (1998). Cooperative R&D and the value of the firm. *Review of Industrial Organization 13*, 425–446.

Xu, M. and C. Zhang (2004). The explanatory power of R&D for the cross-section of stock returns: Japan 1985-2000. *Pacific-Basin Finance Journal 12*(3), 245–269.

Young, A. (1998). Growth without scale effects. *The Journal of Political Economy 106*(1), 41–63.

Zarutskie, R. (2006, September). Evidence on the effects of bank competition on firm borrowing and investment. *Journal of Financial Economics 81*(3), 503–537.

Zhang, A., Y. Zhang, and R. Zhao (2003). A study of the R&D efficiency and productivity of Chinese firms. *Journal of Comparative Economics 31*(3), 444–464.

Samenvatting (Summary in Dutch)

Technologische vooruitgang is een belangrijke bron van economische groei. In veel studies wordt er geconstateerd dat bedrijfsprestaties positief gerelateerd zijn aan innovatie-activtiviteiten zoals onderzoek en ontwikkeling (R&D). Deze activiteiten kunnen resulteren in succesvolle innovaties waardoor bedrijven hun concurrenten voorbij kunnen streven en derhalve hogere winsten realiseren. Onderzoek toont ook aan dat de opbrengsten van R&D investeringen voor de samenleving hoger zijn dan de private opbrengsten van innovatie. De extra opbrengsten van R&D voor de samenleving kunnen worden veroorzaakt doordat bedrijven mogelijkerwijs van elkaar leren tijdens innovatieprocessen waar ze zelf niet direct bij betrokken zijn. De erkenning van de belangrijke rol van technologische vooruitgang voor bedrijfsprestaties en economische groei heeft een zoektocht naar de drijfveren van innovatie ontketend. De zoektocht naar deze drijfveren en het invloedrijke werk van Schumpeter (1942) hebben onder onderzoekers en politici geleid tot een grote interesse in de relatie tussen concurrentie en innovatie. Vanuit het Schumpeteriaanse perspectief is er een positief verband tussen monopoliemacht en innovatie, doordat hoge winsten bedrijven in staat stellen om innovaties intern te financieren en de meest innovatieve mensen in dienst te nemen. Hierdoor wordt een hoge concurrentiedruk gezien als negatief voor het innovatiegedrag van bedrijven. Empirische studies bevestigen een negatief verband tussen concurrentie en innovatie. Desalniettemin, wordt er in enkele empirische studies ook een positief verband gevonden tussen concurrentie en innovatie. Aghion et al. (2001) beargumenteren dat concurrentie een positieve invloed heeft op innovatie, doordat het bedrijven stimuleert om hun concurrenten voorbij te streven. Ook zijn er omgekeerde U relaties gevonden in de literatuur. De verschillende empirische bevindingen zijn afhankelijk van de data en de econometrische specificaties die gebruikt worden.

Het invloedrijke werk van Aghion et al. (2005b) and Aghion and Griffith (2005) heeft de verschillende bevindingen in de theoretische en empirische literatuur samen gebracht door een theoretisch model te ontwikkelen dat zowel een positieve, een negatieve alsook een omgekeerde U relatie tussen concurrentie en innovatie kan verklaren. Als het initiële niveau van concurrentie laag is, zal meer concurrentie een positief effect hebben op het innovatiegedrag van bedrijven, doordat ze geprikkeld worden om hun concurrenten voorbij te streven. Een hogere concurrentiedruk zal pas een negatief effect hebben op innovatie als het initiële niveau van concurrentie hoog is. Dit effect zal dan negatief zijn, doordat meer concurrentie de verwachte winsten zal bedrukken. Ondanks al deze ontwikkelingen in de concurrentie-innovatie literatuur, resteren er nog vele uitdagingen in het kader van het meten van innovatie, het vaststellen van causale verbanden en het rekening houden met de individuele karakteristieken van bedrijven en industrieën.

Volgens Kamien and Schwartz (1982) is het meten van innovatie een van de meest fundamentele problemen in de bestaande literatuur. De conventionele maatstaven zoals patenten, het aantal innovaties, R&D uitgaven en het aantal personeel dat betrokken is bij R&D worden vaak gebruikt in de literatuur. Desondanks hebben deze maatstaven beperkingen. Niet alle innovaties worden namelijk gepatenteerd, het is moeilijk om rekening te houden met de economische waarde van patenten en innovaties vinden ook zonder R&D plaats. Naast deze beperkingen wordt het nut van deze maatstaven in onderzoek beperkt, doordat deze maatstaven in de dienstensector niet altijd beschikbaar zijn. Ondanks dat de dienstensector het belangrijkste deel van de economische activiteiten in ontwikkelde landen omvat, is er door de beperkte beschikbaarheid van innovatiemaatstaven relatief minder onderzoek verricht naar het innovatiegedrag van bedrijven in de dienstensector. Een groot deel van de activiteiten in de dienstensector vindt plaats in de financiële dienstverlening. Onderzoek naar innovatiegedrag van financiële instellingen is belangrijk gegeven de cruciale rol van financiële instellingen in een economie. Innovaties in de financiële dienstverlening faciliteren mogelijk een belangrijke bron van economische groei, namelijk financiële ontwikkeling. Ondanks de belangrijke functie van financiële instellingen moet de relatie tussen concurrentie en innovatie in de financiële dienstverlening nog onderzocht worden. Bestaande studies beperken zich over het algemeen tot een bepaalde technologische ontwikkeling, de focus ligt niet op de relatie tussen concurrentie en innovatie en er wordt vaak geen rekening gehouden met heterogeniteit in het innovatiegedrag van bedrijven. Een andere beperking in veel studies is het onderzoeken van causale verbanden tussen concurrentie en innovatie.

Het meten van innovatie en het rekening houden met heterogeniteit in het innovatiegedrag van bedrijven zijn niet de enige factoren die onderzoek naar het verband tussen concurrentie en innovatie bemoeilijken. Het is algemeen aanvaard dat veranderingen in de marktstructuur en het niveau van concurrentie het innovatiegedrag van bedrijven beïnvloeden. Echter, in de bestaande literatuur is relatief minder onderzoek verricht naar het effect van innovatie op concurrentie. Ook in deze stroming van de literatuur worden er zowel positieve als negatieve effecten van innovatie op concurrentie gevonden. Een beperking van de huidige studies is dat bij het onderzoeken van het effect van innovatie op concurrentie, de focus slechts ligt op de factoren die gebruikt worden om innovaties te produceren of op de uiteindelijke, meetbare innovaties. Beide aspecten hangen echter met elkaar samen en bevatten afzonderlijk informatie over het niveau van concurrentie. De verschillende bevindingen in de huidige literatuur zijn daarom afhankelijk van de maatstaf die gehanteerd wordt bij het meten van innovatie. Meer inzichten in de effectiviteit van innovatie voor de concurrentiepositie van bedrijven is niet alleen belangrijk voor bedrijven zelf, maar ook voor overheden aangezien innovatiebeleid aangewend kan worden om het niveau van concurrentie op een markt te beïnvloeden.

Ondanks dat politici en onderzoekers erkennen dat de relatie tussen concurrentie en innovatie belangrijk is, suggereren de verschillende bevindingen in de bestaande literatuur dat potentiële kanalen en de relatie tussen de factoren die gebruikt worden om innovaties te produceren en de resulterende innovaties, nader onderzocht dienen te worden.

In dit proefschrift wordt een nieuwe innovatiemaatstaf geïntroduceerd en de causale verbanden tussen concurrentie en innovatie vastgesteld. Ten eerste wordt er een nieuwe innovatiemaatstaf geïntroduceerd door de relatieve technologische achterstand van banken (ten opzichte van elkaar) te onderzoeken. Deze nieuwe innovatiemaatstaf maakt het mogelijk om de relatie tussen concurrentie en innovatie te onderzoeken in sectoren waar conventionele innovatiemaatstaven niet ter beschikking zijn. Ten tweede wordt in dit proefschrift de relatie tussen concurrentie en innovatie nader onderzocht door belangrijke marktontwikkelingen in de bankensector, zoals de consolidatietrend gedurende de afgelopen twee decennia, te onderzoeken. Specifiek wordt het effect van schaalvoordelen op het verband tussen concurrentie en innovatie geanalyseerd in het kader van de impact die de consolidatietrend heeft gehad op de marktstructuur en concurrentie. Ten derde bevat dit proefschrift een analyse waarin het effect van innovatie op concurrentie onderzocht wordt op basis van Nederlandse innovatiestatistieken. In deze analyse wordt onderscheid gemaakt tussen factoren die gebruikt worden in het innovatieproces en de resulterende innovaties. Aangezien deze concepten nauw met elkaar samenhangen, richt dit proefschrift zich ten vierde op de relatie tussen de factoren die gebruikt worden in het innovatieproces en de resulterende innovaties. In deze relatie wordt ook de efficiëntie in de productie van innovaties nader onderzocht.

In hoofdstuk 2 wordt een nieuwe innovatiemaatstaf geïntroduceerd om de relatie tussen concurrentie en innovatie in de bankensector van de V.S. te onderzoeken. De nieuwe innovatiemaatstaf wordt verkregen door stochastische grensmodellen te schatten en op basis hiervan de jaarlijkse kostengrenzen en wereldgrens te berekenen. De wereldgrens weergeeft de beste potentiële technologie die beschikbaar is voor banken. De afstand tot deze wereldgrens weergeeft de technologische achterstand die een bank ondervindt. Innovaties stellen banken in staat om hun technologische achterstand te verkleinen. De nieuwe innovatiemaatstaf wordt gebruikt om het model van Aghion et al. (2005b) op bankniveau te schatten voor de V.S. in de periode 1984-2004. Gedurende deze periode was er sprake van een consoliderende markt met enkele regulatieveranderingen. De effecten van deze marktontwikkelingen op concurrentie en innovatie worden onderzocht. De resultaten in dit hoofdstuk wijzen op een omgekeerde U relatie tussen concurrentie en innovatie in de bankensector in de V.S.. Meer concurrentie heeft een positief effect op innovatie bij een lager niveau van concurrentie en een negatief effect bij een hoger niveau van concurrentie. De empirische analyse toont ook aan dat de winstmarges van banken gestegen zijn tijdens de consolidatietrend. Ook de deregulatie waardoor het voor banken mogelijk werd gemaakt om de staatsgrenzen te verleggen, heeft een negatief effect gehad op concurrentie en indirect op het innovatiegedrag van banken.

In hoofdstuk 3 worden de theoretische en empirische analyses van hoofdstuk 2 en Aghion et al. (2005b) uitgebreid om het verband tussen concurrentie en innovatie in een consoliderende markt nader te onderzoeken. Specifiek staat de vraag centraal of banken te groot geworden zijn om te innoveren. In het model van Aghion et al. (2005b) wordt er verondersteld dat de kosten per eenheid product onafhankelijk zijn van de geproduceerde hoeveelheid. Het model in hoofdstuk 3 gaat uit van een gemiddelde kostencurve met een U vorm, zoals vaak geconstateerd wordt in de bankensector. Zodoende ondervinden bedrijven in het model schaalvoordelen of schaalnadelen. In het theoretische model in hoofdstuk 3 worden algemene condities afgeleid waaruit opgemaakt kan worden onder welke omstandigheden het innovatiegedrag van bedrijven met schaalnadelen gevoeliger of minder gevoelig zijn voor veranderingen in concurrentie. In hoofdstuk 3 wordt dezelfde innovatiemaatstaf gebruikt als in hoofdstuk 2, namelijk de relatieve technologische achterstand van banken. De empirische resultaten in hoofdstuk 3 laten zien dat banken gemiddeld over de tijd heen steeds meer schaalnadelen ervaren, terwijl de winstmarges gemiddeld gestegen zijn. De resultaten laten ook zien dat er sprake is van een omgekeerde U relatie tussen concurrentie en innovatie en dat het innovatiegedrag van grote banken gevoeliger is voor veranderingen in concurrentie.

De analyses in hoofdstuk 2 en 3 verschaffen meer inzicht in het verband tussen concurrentie en innovatie in de bankensector van de V.S. en marktontwikkelingen zoals consolidatie en deregulatie. De meeste banken hebben hogere winstmarges dan de optimale winstmarge die innovatie maximaal bevorderd. De resultaten impliceren dat ontwikkelingen in de toekomst die het niveau van concurrentie verder verlagen, de innovatieprikkels van banken negatief beïnvloeden. Met name wordt het innovatiegedrag van grote banken meer dan proportioneel negatief beïnvloed. Desalniettemin impliceren de resultaten in hoofdstuk 2 en 3 ook dat beleidsmatige veranderingen die voor meer concurrentie tussen banken zorgen, tot meer innovaties zullen leiden. In een geconsolideerde markt zullen de grote banken hier relatief sterker op reageren.

In hoofdstuk 4 ligt de nadruk op het effect van innovatie op concurrentie. Specifiek wordt in dit hoofdstuk onderzocht of factoren die gebruikt worden bij het produceren van innovaties een ander effect op concurrentie zullen hebben dan hetgeen geproduceerd wordt in het innovatieproces. In de bestaande literatuur worden deze dimensies van innovatie voornamelijk afzonderlijk van elkaar onderzocht. In hoofdstuk 4 wordt er gebruik gemaakt van een theoretisch model waarin collusie tussen bedrijven gemodelleerd wordt. Een belangrijke factor in de productie van innovaties is de opgedane ervaring tijdens innovatieprocessen in het heden en verleden. In het model wordt er verondersteld dat meer ervaring de kans zal verhogen dat er een innovatie geproduceerd zal worden en collusie tussen bedrijven zal verhinderen. Ondanks dat een verhoogde kans op innovatie tot meer concurrentie zal leiden, zal de realisatie van een innovatie uiteindelijk concurrentie op een markt verminderen aangezien innovatieve bedrijven dominanter worden op een markt. Innovatiestatistieken van Nederlandse bedrijven over de periode 1994-2004 worden gebruikt om het effect van deze dimensies van innovatie op concurrentie empirisch te onderzoeken. De bevindingen tonen aan dat gerealiseerde innovaties tot minder concurrentie leiden, terwijl meer ervaring met het produceren van innovaties zal leiden tot meer concurrentie. Tevens is het effect van een gerealiseerde innovatie sterker dan het effect van ervaring met het produceren van innovaties. Het onderzoeken van het effect van deze dimensies van innovatie op concurrentie is belangrijk om de relatieve grootte van de effecten vast te stellen en vanwege het feit dat deze dimensies met elkaar samenhangen. Vanuit een beleidsperspectief impliceren de resultaten dat R&D beleid en patentenbeleid niet alleen innovatiegedrag van bedrijven beïnvloeden, maar ook de concurrentiepositie van bedrijven. Tevens kunnen de verschillende dimensies van innovatie voor feedback effecten zorgen, waardoor initiële concurrentieprikkels versterkt of juist verzwakt kunnen worden.

Hoofdstuk 5 bevat een gedetailleerde analyse van de relatie tussen deze dimensies van innovatie door het productieproces van innovaties te onderzoeken. Met name wordt de relatie tussen factoren die gebruikt worden bij het produceren van innovaties en hetgeen wat er in het innovatieproces geproduceerd wordt onderzocht. Tevens bevat hoofdstuk 5 een analyse van de mate van inefficiëntie in het innovatieproces en de determinanten hiervan. Nederlandse innovatiestatistieken en stochastische grensmodellen worden gebruikt om een innovatie productiefunctie te schatten en inefficiëntie te identificeren. Tevens worden de gebruikte factoren in het innovatieproces en inefficiëntie gerelateerd aan de samenwerking tussen bedrijven, samenwerking met andere instellingen, subsidie van de overheid, concurrentie en bedrijfsomvang. De resultaten in hoofdstuk 5 laten zien dat de innovatie productiefunctie gekenmerkt wordt door afnemende schaalopbrengsten en dat er een inefficiënte innovatieproces plaatsvindt. Inefficiëntie blijkt een belangrijke factor te zijn in het innovatieproces aangezien het ongeveer 63% van de productiviteitsverschillen verklaart. Meer concurrentie en een grotere bedrijfsomvang zijn positief gerelateerd aan de productie van innovaties. Grote bedrijven blijken tevens minder efficiënt innovaties te produceren. De bevindingen laten ook zien dat subsidies van de overheid in het verleden de efficiëntie in het innovatieproces in de toekomst bevorderen. Een belangrijke beleidsconclusie van hoofdstuk 5 is dat als subsidies de efficiëntie verhogen, subsidies een effectieve hulpmiddel blijken te zijn om de productie van innovaties te stimuleren.

Samenvattend, wordt er een nieuwe innovatiemaatstaf in hoofdstuk 2 geïntroduceerd om de relatie tussen concurrentie en innovatie in de Amerikaanse bankensector te onderzoeken, aangezien conventionele maatstaven niet beschikbaar zijn. In hoofdstuk 3 wordt het theoretisch model van Aghion et al. (2005b) aangepast om de effecten van de consolidatietrend nader te onderzoeken door het effect van schaalnadelen op het verband tussen concurrentie en innovatie te analyseren. De belangrijkste conclusies van hoofdstuk 2 en 3 zijn dat er een omgekeerde U relatie is tussen concurrentie en innovatie in de Amerikaanse bankensector, dat consolidatie en deregulering tot minder concurrentie en innovatie hebben geleid en dat voornamelijk het innovatiegedrag van grote bedrijven relatief gevoeliger is voor veranderingen in concurrentie. In tegenstelling tot hoofdstuk 2 en 3, wordt in hoofdstuk 4 het effect van innovatie op concurrentie onderzocht. Het gebruik van meer input factoren bij de productie van innovaties leiden tot meer concurrentie, terwijl gerealiseerde innovaties voor minder concurrentie zorgen. Door de samenhang van deze dimensies is het noodzakelijk dat de effecten van deze dimensies op concurrentie in een raamwerk onderzocht worden. Hoofdstuk 5 gaat nader in op deze samenhang en toont aan dat er een positief verband is tussen het gebruik van factoren bij de productie van innovaties en het realiseren van innovaties. Tevens verloopt het innovatieproces inefficiënt en zal het aanpassen van inefficiënte innovatieprocessen de productie van innovaties verhogen.

Curriculum Vitae

I, Ryan Christopher Richard van Lamoen, was born in Willemstad, Curacao on October 21, 1983. From 1996 until 2002, I attended the Atheneum at Cals College in Nieuwegein. From 2002 until 2005, I enrolled in the Economics bachelor program at the Utrecht University School of Economics and finished the tracks Financial-Monetary Economics, International Economics, Business & Management and Labor & Organization. Subsequently, I enrolled in the master program International Economics and Business at the Utrecht University School of Economics and followed the track International Managerial Economics. I wrote my master thesis and research paper entitled "Go with the Flow: Theory and Practice of Underwriting Cycles in Insurance Sectors" during an internship at the Quantitative Risk Management department of De Nederlandsche Bank. After graduating with honors in 2006 with a Master's degree in International Economics and Business, I started with my PhD research at the Utrecht University School of Economics on the relationship between competition and innovation. In 2010, I was awarded the "Best PhD Student Paper Award" at the 2010 Conference on "Global Trends in the Efficiency & Risk Management of Financial Services" in Chania, Greece. At present, I work as a Model Validator at the Group Risk Management department of SNS REAAL.

TKI Dissertation Series

USE 001 **Bastian Westbrock** (2010): *Inter-firm Networks: Economic and Sociological Perspectives.*

USE 002 **Yi Zhang** (2011): *Institutions and International Investments: Evidence from China and Other Emerging Markets.*