ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
GÖTEBORG UNIVERSITY
168

Essays on Performance and Growth in Swedish Banking

Pål Sjöberg

ISBN 91-85169-27-7
ISBN 978-91-85169-27-6
ISSN 1651-4289 print
ISSN 1651-4297 online

GÖTEBORG UNIVERSITY

To my father and late mother,
Clas and Birgitta

Abstract

This thesis deals with performance and growth in the Swedish banking sector, in an era following important changes such as the globalisation of financial markets, the harmonisation of legislation (e.g. the EU banking directives) and the implementation of new technology, such as Internet banking and other electronic delivery channels. The thesis consists of an introductory chapter and three self-contained papers.

Paper 1: Market power and performance in Swedish Banking

This paper analyses the degree of competition in the Swedish banking market, over the period 1996-2002. A structural simultaneous-equation model, which includes a conduct parameter, is estimated. The results indicate that the average bank's conduct was more competitive than Cournot behaviour, although not perfectly competitive. The average Lerner index obtained equals 22%. Furthermore, the results indicate that competition among commercial banks was significantly more intense than among savings banks, despite the formers' much larger size. This finding may suggest that commercial banks operate in a more contestable environment. Finally, the results show that banks operating more than one office experienced significantly better performance, in terms of higher price-cost margins, than one-office banks (unit banks).

Paper 2: Competition in Swedish Local Banking Markets

In contrast to urban and metropolitan bank customers, rural bank customers still rely on the physical network (branches) as the prime access channel. This means that high customer loyalty and entry barriers can be expected to prevail in rural banking. Against this background, the paper analyses the degree of competition in rural banking markets, using a variation of the Bresnahan and Reiss entry model. According to the results, entry thresholds increase more than proportionately with each additional entry, suggesting that profit margins shrink as a result of new entry. The resulting pro-competitive effect is most pronounced in markets with a relatively few number of competitors. Finally, the results suggest that a greater share of "multi-market banks" in a given market promotes local competition – a result which parallels a number of international studies.

Paper 3: A Dynamic Analysis of Firm Growth in Swedish Banking

This paper examines firm growth dynamics in the new banking environment, by testing the validity of Gibrat's Law of Proportionate Effect on Swedish data. The point of departure in the paper is the expectation that large banks should be able to more fully exploit scale and scope economies associated with technological innovations such as internet banking, than smaller banks, and therefore grow faster. Using a panel of 79 Swedish banks over the period 1995-2002, I find no empirical evidence that large banks grew faster, nor any significant evidence that firm sizes were mean-reverting. Hence the Law was not rejected. However, growth was not entirely random, as banks with a more diversified revenue mix experienced significantly higher growth rates than their less diversified counterparts.

JEL Classification: C23; C30; G21; L11; L13.

*Keywords***:** Swedish banking, degree of competition, Lerner indexes, entry barriers, contestability, commercial banks, savings banks, branch banks, unit banks, GMM estimation, dynamic panel data model, firm growth, Gibrat's law of proportionate effect, technological change, local banking markets, entry thresholds, ordered probit model, Poisson model.

Preface

In the course of writing this thesis, I have benefited from several people. First of all, I would like to express my deepest gratitude to my supervisor, Professor Lennart Hjalmarsson for his guidance, support, and everlasting enthusiasm and patience throughout the years of this work. Thank you!

I would also like to express my sincere thanks to other colleagues at the department who have had an impact on this work. In particular, I am indebted to Associate Professor Dick Durevall, Professor Lennart Flood, Professor Rune Stenbacka and Associate Professor Johan Lönnroth for many helpful comments on earlier drafts of the dissertation papers. A special thanks goes to Dick Durevall, with whom I had many inspiring and fruitful discussions on different aspects of the work.

The administrative staff at the department deserves a special thanks for excellent help with practical matters during my thesis work. Thank you so much, Elizabeth Földi, Eva Jonason, Gunilla Leander, Ulla Mellgren, Eva-Lena Neth-Johansson and Katarina Renström.

During my work with the thesis, I was fortunate to receive financial support from Göteborg University, Konkurrensverket and Jan Wallander's och Tom Hedelius' Foundation. The author gratefully acknowledges this support.

I am very grateful to the comments provided by Professor Mattias Ganslandt, who was the opponent of my licentiate thesis. His comments have been very helpful to improve the first and third chapters of the thesis. Moreover, I owe Professor Mats Bergman, a great deal for helping me with the first paper in the thesis, as well as Professor Per-Olof Bjuggren, for providing useful comments on an earlier draft of the second thesis paper.

I wish to thank my fellow colleagues and new friends, who made my graduate studies such a pleasant experience. I thank Fredrik Andersson, Daniel Deng, Jorge Garcia, Alexis Palma, and Elias Tsakas for great company during the course of writing my thesis. My special thanks go to Jens Madsen and Klas Sandén, with whom I shared many enjoyable moments. Thank you, Klas and Jens, for all the enjoyable lunchtime conversations!

My final words of gratitude go to my family. I sincerely thank my extended family – brothers, sisters, in-laws and my many nieces and nephews for their emotional support. Above all, however, I am forever indebted to my beloved parents, Clas and Birgitta, for their everlasting support and encouragement. Very tragically and unexpectedly, however, my mother Birgitta passed away before my journey was completed. Wherever she is now, I know she is happy to see that I finally have accomplished my goal.

Despite all the helpful comments and suggestions that I have received during the course of writing this thesis, errors may still remain. Whatever errors remain are entirely mine.

Pål Sjöberg

Göteborg, October 2007.

Contents

Introduction

The Swedish banking sector has experienced fundamental changes since the beginning of the ‘90s. In 1991-1992, Sweden experienced a severe banking crisis. In the mid ‘90s, when confidence in the Swedish banking sector was restored, and the new banking environment was in place, the banking sector witnessed the entry by new players such as foreign banks operating through branches or subsidiaries. Moreover, since the delineations between the banking and insurance sectors became blurred in the mid ‘90s, several insurance companies opened niche banks (internet banks).

Important changes that have contributed to a new banking environment are the globalisation of financial markets, the harmonization of relevant legislation (e.g. the EU banking directives) aimed at reducing cross-country entry barriers, and the implementation of Internet banking and other electronic delivery channels.

By reducing entry barriers, the new IT-based technology and the harmonized legislation are expected to have intensified competition in the banking sector. As mentioned above, there was a response, in terms of new entry, to these changes. Besides the aforementioned types of banks, companies with large customer bases, such as ICA and COOP, have subsequently opened or attempted to open banks.

However, the expected increase in competitive conditions in the banking sector is not attributed to actual entry only. To the extent that the aforementioned changes have reduced entry barriers, potential competition should come into play. In fact, the issue of potential competition appears to be an issue of perhaps more relevance than ever to the banking sector (ECON Report 25, 2007). If potential competition is indeed effective, the market should exhibit characteristics of contestability.

The assessment of competitive intensity in the new, presumably more contestable banking environment is the subject of the first two papers of the thesis. The first paper evaluates overall competitive conditions in the banking industry over the period 1996-2002, using a structural econometric model founded on microeconomic theory.

I find that the behaviour of the average bank was quite competitive, reflecting that the degree of market power possessed by the average firm was quite low. It should be noted, however, that since banks are suppliers of a broad range of financial services, price-margins may differ across the various product markets. Such differences could reflect differences in competitive conditions among the markets or that incumbent banks erect strategic entry barriers through predatory pricing in certain segments, while extracting rents in other. On average however, the intense of competition is high in view of the high degree of concentration of the banking sector. Thus, the results do not lend much support to hypotheses that predict a distinct market concentration- market power relationship, such as the Cournot model. Given that Cournot conduct would imply an average Lerner index of 35% while the estimated average Lerner index is 22%, the hypothesis of Cournot conduct could be rejected. The obtained value of the Lerner index is comparable in magnitude to that reported in Maudos and Nagore (2005), which equals 20%.

Furthermore, the results indicate that commercial banks are significantly more competitive than savings banks, on average. Put differently, savings banks enjoy higher price-cost margins than commercial banks, on average. This result reflects that commercial banks operate in a more competitive environment than savings banks. In view of the fact that the commercial banks are much larger (and fewer in number) than the savings banks, it is again clear that the results do not lend empirical support to hypotheses that infer conduct based on concentration characteristics. By contrast, the result may indicate that the competitive environment facing commercial banks exhibit characteristics of contestability.[1]

In view of the result that savings banks appear to perform in a less competitive environment, the objective of the second paper is to analyse the intense and mode of competition in markets where savings banks are present and often predominant, i.e. local (rural) markets. In these markets savings banks, besides the largest nationwide banks, supply services through a network of branches, as well as online.

According to a recent survey (Svenskt Kvalitetsindex, 2006), rural bank customers still rely on the physical network (branches) as the prime distribution channel for conducting retail banking services such as savings and lending, while for urban citizens online banking has become the prime channel. *A priori*, this finding suggests that anti-competitive entry barriers associated with traditional branch banking remain an important issue in rural banking.

[1] Recent international studies tend to show that large banks operate in a more competitive environment (national and international markets) than small banks, which operate in local markets. See e.g. Bikker and Haaf (2002).

In the paper, a modified version of the Bresnahan-Reiss game theoretic entry model is employed.[2] The model is a two-stage game where competitors in the first stage simultaneously decide on whether or not to enter a particular local market. Conditional on entry, they in the second stage participate in a price game. As usual, I focus on the entry stage of game, using an ordered probit model to estimate entry thresholds for retail banking service providers.[3] The sample covers the period 1998-2002, and 97 rural markets, proxied by local labour market areas.

The results suggest that profit margins in the retail branch banking industry must to be quite high in the most concentrated markets, despite the relatively larger presence of nationwide banks in these markets. However, margins fall substantially with each additional entrant, suggesting a clear relationship between the degree of concentration and the market power possessed by each supplier. The conclusions that can be drawn from these results are threefold. First, there is no empirical support for the hypothesis of contestability at the local level. If local markets were contestable, no distinct concentration-profit margins relationship would be obtained. Second, there is no evidence that banks collude as a cartel, since profit margins always fall with each additional entrant. Third, the gradually decreasing pattern of profit margins as the number of competitors increase is consistent with the assumption made in the paper that retail branch banks offer relatively homogenous services.

To my knowledge, no similar study based on Swedish data has previously been undertaken that the results could be compared with.[4] Neither are the results in Paper 2 directly comparable to those of Paper 1, since the model employed in Paper 2 identifies *changes* in competitive conditions with each additional entry, not the degree of competition itself. We can, however, conclude that there is no inconsistency between the results. As mentioned above, the results from Paper 1 show that savings banks operate in a significantly less competitive environment (local markets) than commercial banks, despite the fact that the latter in general are much larger. It was suggested above that the competitive environment facing commercial banks might exhibit contestable characteristics. The results from Paper 2 indicate that this is not the case in local markets, where savings banks are predominant. Also, the result from

[2] See Bresnahan and Reiss (1991).
[3] Entry threshold= the minimum market size necessary to support a given number of firms.
[4] Cetorelli (2002) investigates competitive conditions in US local (rural) banking markets, using a similar model. As in the present case, the results indicate that profit margins shrink as a result of entry. Likewise, the pro-competitive effect of entry is most pronounced in the most concentrated markets.

Paper 1 that savings banks operate in a less competitive environment, i.e. average price-cost margins are relatively high in local markets, is not contradicted by the results in Paper 2.

The third paper of the thesis examines firm growth dynamics in the new banking environment. More specifically, I test if firm growth rates obey Gibrat's Law of Proportionate Effect (LPE). Under the LPE, firm growth rate is independent of firm size (and previous growth performance). As well known, the implication of the LPE is a firm size distribution which becomes increasingly skewed over time (and dominated by a small number of large firms), and eventually converges to the lognormal distribution.

Due to the fact that the empirical firm size distributions in many industries resemble the lognormal distribution, the LPE has drawn a substantial amount of attention over the years. Essentially all the empirical literature testing for the LPE has concerned manufacturing industries. Within manufacturing, the stylized facts of the empirical literature are that (1) firm sizes are mean-reverting, which contradicts to the LPE, while (2) for sub-samples of large and well-established firms, the LPE tends to be confirmed (Gibrat's legacy).[5]

By contrast, the limited number of tests of Gibrat's law based on banking data does not suggest evidence of mean-reversion. Rather, the evidence suggests either a non-relationship between size and growth, in accordance with the LPE, or a weak positive relationship.[6] In any case, the implication is that concentration in the banking sector will continue to increase.

The main objective of Paper 3 is to contribute to the understanding of firm growth dynamics in the new banking environment, by testing for the LPE on Swedish data. The point of departure in the paper is the expectation that large (nationwide) banks should be able to more fully exploit scale and scope economies associated with technological innovations such as internet banking, than smaller banks, and therefore grow faster.

Using dynamic panel data techniques (GMM estimation), I estimate a multivariate firm-growth equation in which I control for various bank-specific determinants, such as profitability, efficiency and diversification. The period covered is 1995-2002. Based on a large cross-section of Swedish banks, I find no empirical

---

[5] See Sutton (1997).
[6] See Goddard *et al.* (2001) and Goddard *et al*. (2004).

evidence that large banks grew faster, nor any significant evidence that firm sizes were mean-reverting. Hence, the LPE could not be rejected. However, growth is not entirely random, as banks with a broader product range experienced significantly higher growth rates than less diversified banks.

Bibliography

Bikker, J.A. and Haaf, K. (2002), ”Competition, Concentration and their relationship: An Empirical Analysis of the Banking Industry”, *Journal of Banking and Finance*, vol. 26, pp. 2191-2214.

Bresnahan, T. and Reiss, P. (1991), “Entry and Competition in Concentrated Markets”, *Journal of Political Economy*, vol. 99, pp. 977-1009.

Cetorelli, N. (2002), “Entry and Competition in Highly Concentrated Banking Markets”, *Economic Perspectives*, Q4, pp. 18-27.

ECON Report (2007),“Konkurrensen på Bankmarknaden” http://www.econ.se

Goddard, J.A., Molyneux, P. and Wilson, J.O.S. (2001), *European Banking – Efficiency, Technology and Growth*, John Wiley & Sons, Ltd, England.

Goddard, J.A., Molyneux, P. and Wilson, J. (2004), ”Dynamics of Growth and Profitability in Banking”, *Journal of Money, Credit, and Banking*, vol. 36, pp. 1070-1090.

Maudos, J. and Nagore, A. (2005), ”Explaining Market Power Differences in Banking: A Cross-country Study”, University of Valencia, Working Paper

Sutton, J. (1997), “Gibrat’s Legacy”, *Journal of Economic Literature*, vol. 35, pp. 40-59.

Svenskt Kvalitetsindex (2006), 2006-10-02.

Market Power and Performance in Swedish Banking

By

Pål Sjöberg[*]

Department of Economics and Statistics
Göteborg University

Abstract

The objective of this study is to empirically assess the degree of competition in the Swedish banking market, over the period 1996-2002. For this purpose, a structural simultaneous-equation model, in which a conduct parameter is embedded, is estimated, using the GMM and SUR estimation procedures. The results indicate that the average bank's behaviour was more competitive than what the Cournot model would imply, although far from perfectly competitive. Furthermore, the results suggest that competition among commercial banks was significantly more intense than competition among savings banks. Put differently, commercial banks were operating in a more competitive environment than savings banks. This is consistent with the hypothesis of contestability – commercial banks operate in business segments where entry barriers are comparatively low. Finally, the results suggest that banks operating more than one office experienced significantly better performance, in terms of price-cost margins, than one-office banks (unit banks).

JEL Classification: C30; G21; L13.

Keywords: Swedish banking, degree of competition, Lerner indexes, contestability, commercial banks, savings banks, branch banks, unit banks, GMM estimation.

[*] I thank Mats Bergman, Dick Durevall, Mattias Ganslandt, Lennart Hjalmarsson, Johan Lönnroth and Rune Stenbacka for valuable comments and improvement suggestions. Financial support from the Swedish Competition Authority is gratefully acknowledged.

1. Introduction

In many respects, the nature of banking has changed over the past 20 years. The new banking environment is a result of, *inter alia*, the deregulation of capital markets, consolidation, the implementation of harmonized legislation (e.g. the EU banking directives) aimed at reducing cross-country entry barriers, and technological changes. Technological innovations such as internet banking (which was first launched in Sweden in 1995) and other new-technology delivery systems have opened up new delivery channels for banking services. As a result, banks have been able to reduce transaction costs substantially, and exploit economies of scale associated with transaction processing. Partly due to the extensive use of electronic banking transactions, Swedish banks are among the most efficient and productive in Europe. Thus, during the period under study (1996-2002), Swedish banks experienced an average annual increase in productivity as high as 4.6%. Furthermore, the Swedish banks are world leaders in the supply of internet banking services.[7]

A priori, changes such as the new legislation and the implementation of new technology can be expected to have intensified actual as well as potential competition in the banking sector, by reducing entry barriers. Indeed, the Swedish banking industry experienced entry by foreign banks in certain business segments, such as the corporate market, during the period under study. Moreover, non-banking domestic companies opened internet banks around 1995. However, the expected increase in competition is not attributed to actual entry only. By reducing entry barriers, the aforementioned changes could be expected to have increased potential competition as well. If this conjecture is correct, the industry (or part of it) should exhibit characteristics of contestability.[8]

To my knowledge, no recent and comprehensive investigation, aimed at empirically assessing competitive conditions in the new Swedish banking environment, has been undertaken. The main purpose of the present paper is thus to fill this void by estimating the degree of competition in this industry, using a structural empirical model in which a parameter measuring banks' competitive behaviour (conduct) is embedded. An additional objective of the paper is to

[7] See ECON Report 25 (2007).

[8] A contestable market has low barriers to entry and exit. If the market is perfectly contestable, entry and exit are totally costless, i.e. there are no sunk entry costs. A high degree of contestability (potential competition) may render a market perfectly competitive regardless if the industry is highly concentrated or not. See Baumol *et al*. (1982).

investigate if certain types of banks differ in terms of competitive behaviour, and thus possess different degrees of market power.

The empirical model is a modified version of the Bresnahan and Lau (1982) model. A simultaneous equation model, comprised by a demand equation, a cost equation and a supply equation, is estimated. The supply equation is derived assuming profit maximizing behaviour among banks. The system is estimated using the GMM and SUR estimation methods. The main result is that the average bank's behaviour is more competitive than the Cournot model would predict.

The rest of the paper is organized as follows. Chapter 2 briefly describes the recent evolution of the market, while Chapter 3 gives an orientation about the literature in this field. In Chapter 4, the methodology is thoroughly described. Chapter 5 is devoted to a discussion and analysis of the data, while Chapter 6 reports and interprets the results. Finally, Chapter 7 concludes.

2. Background

During the last two decades, the Swedish financial market has witnessed a fundamental transformation process, resulting in a drastic vitalization of its performance. Before the financial crisis, which plagued the country in the beginning of the 90's, the market saw an increase in the number of players, as a response to the important deregulation step, undertaken in 1986, implying the abolition of the prohibition of foreign bank participation.[9] Subsequently, in connection to the severe banking crisis in 1991-1992, a substantial decrease in the number of banks was observed. However, since the financial market was restored around 1993, the number of banks again started to rise. Once again, the market witnessed an increase of foreign banks into segments such as the corporate market, a development triggered by the decision to lift the prohibition of opening branches in Sweden[10]. Moreover, domestic niche banks (internet banks) such as Ikanobanken, Länsförsäkringar Bank and Skandiabanken appeared on the scene. These banks were all founded in the mid '90s. However, during the period under study (1996-2002) the total number of banks has been fairly stable around 125.

---

[9] However, many of these establishments actually failed, since, as it turned out, they had focused on the wrong business segment. (Lybeck, 2000)

[10] This step was taken in 1990. The former legislative change, in force from 1986, allowed foreign banks to open subsidiaries only.

Another apparent feature of the transformation process has been a general tendency of broadening the scope of the business mix. Intensified competition faced in the traditional intermediation business has forced banks to rethink their strategies. While the vitalized capital markets certainly accounts for part of the increased pressure put on the banking sector, there have also been steps taken, of legislative nature, with the aim of increasing competition in the financial market. Important, international steps, aimed towards harmonization and increased competition, has been undertaken. On the global arena, the Bank for International Settlements (BIS)[11] has influenced the way international oriented banks perform, through the Basel Capital Accord (1988). The new accord, Basel II, scheduled for implementation in 2007, is planned to be compulsory through a new EU directive. Today, though, the Second Banking Coordination Directive (1988) is valid. In Sweden it came into force through the EES settlement (1993). Undoubtedly, this directive can be regarded as the cornerstone of the new legislation era. In this directive, the principle of the single market licence allowing banks and other credit banks to set up branches and offer services throughout the Community is established, and it contains a list of banking services that can be provided in all the Member States on the basis of such a licence[12].

Besides the contribution of deregulation, other important driving forces behind the changed conditions in banking can be identified. As mentioned in Chapter 1, the implementation of new technology has profoundly changed the way banking is conducted, and increased efficiency in certain product segments. The launch of the new pension system and a changed demographic structure are other factors that may have contributed to the changed conditions. To cope with the new conditions, incumbents have been forced to adjust their strategies. The largest commercial banks, in particular, have increasingly diversified into new fee- and commission based segments.

3. Related Literature

Traditionally, the degree of competition in a market has been inferred according to the structure-conduct-performance paradigm (SCPP), developed by Bain (1951). Studies relying on this paradigm consider profitability or price as an endogenous variable,

[11] In essence, the BIS is a central bank for central banks. It fosters cooperation among central banks as well as other international financial banks. See http://www.bis.org.
[12] http://europa.eu.int/scadplus/leg/en/lvb/l24002.htm

which are related to market structure characteristics, assumed exogenous. According to the SCPP, a positive relation between structure and performance are expected, because firms in more concentrated markets will earn higher profits (for collusive or monopolistic reasons, the SCPP is not explicit on the issue of conduct, which is one of its weaknesses) than firms operating in less concentrated markets, irrespective of the efficiency performance of individual firms. The empirical support for the SCPP within the banking context is mixed. A couple of influential papers all suggest a positive relationship.[13] By contrast, Jackson (1992), who reviews the results obtained in Berger & Hannan (1989), suggests that the market structure-market power relationship might not be monotonic, in contradiction to the SCPP. Indeed, he shows that for high levels of concentration, a further increase in concentration actually implies less anticompetitive behaviour. In a later study concerning the rigidity of deposit rates, Jackson (1997) again obtains results supporting a non-monotonic structure-market power relationship. Finally, the studies by Rhoades (1995) and Hannan (1997) also cast doubts on the robustness of the concentration-market power relationship.

The SCPP approach has also come under attack by proponents of the so-called efficiency structure hypothesis (EFS).[14] According to the EFS, a positive market structure-market power relationship will emerge as a result of superior efficiency by particular firms, which as a result of their superior performance will gain in market share, implying increased market concentration. Smirlock (1985) performs a direct test between the two competing hypotheses, by running a regression of a performance measure (profitability) on market concentration as well as a variable capturing individual market share (as well as control variables). By finding a positive relationship between the market share variable and profitability, while no relationship between the overall concentration measure and profitability, he concludes that the results are in favour of the EFS (An indication of the SCPP would require the opposite).

Alternative theories suggest conditions under which decisions on pricing and output might be independent of the market structure. One such theory is that of trigger price strategies (Friedman, 1971). This theory shows that collusive behaviour may be sustained among arbitrarily many firms. As another example, the theory of

[13] See Berger & Hannan (1989), Berger & Hannan (1991) and Neumark & Sharpe (1992).
[14] See e.g. Demsetz (1973), Brozen (1982), Smirlock (1985) and Evanoff & Fortier (1988).

contestability (cf. Chapter 1) argues that if entry barriers are non-existent or easily forced, incumbents will be forced to price competitively in order to prevent entry. Tests of contestability have typically relied on estimation of a reduced form revenue function, where the comparative statics of the effect of changes in input prices upon revenue in equilibrium are analyzed (the Panzar-Ross test).[15] This theory has in fact been given a great deal of attention recently, not least within a European banking context, which is understandable in view of the changed market conditions as described in Chapters 1 and 2 above. Empirical applications of the Panzar-Ross methodology to European banking include Molyneux *et al.* (1994), Coccorese (1998), Bikker & Groenevald (2000), DeBandt & Davis (2000), and Bikker & Haaf (2002). Although these tests quite consistently indicate increased characteristics of contestability in the banking markets investigated, the test itself is not reliable unless the sample is in long-run equilibrium. Thus, if the sample is not in long-run equilibrium, the Panzar-Ross index of competitive conditions may exhibit a downward bias towards the collusive oligopoly equilibrium.[16] Furthermore, as a direct test of conduct and performance, where the comparative statics of a profit-maximizing equilibrium is analyzed, the Panzar-Ross test is weak compared to the Bresnahan and Lau test employed in the present paper.[17] Shaffer (2001) argues that the Bresnahan and Lau test has greater ability to econometrically identify the conduct parameter, and to map it into specific oligopoly solution concepts.[18]

With regard to Sweden, Shaffer (2001) obtains evidence of contestability in the commercial banking over the decade preceding the financial crisis, i.e. 1979-1991. By contrast, Oxenstierna (2000), who tests for market power in the Swedish banking oligopoly (composed of the nationwide banks), over the period 1989-1997, obtains evidence of substantial market power in the intermediation margin business, especially on the deposit side.

---

[15] This is the so-called Panzar-Ross test. See Panzar & Rosse (1987)
[16] See Shaffer (1983) and Molyneux *et al*. (1994).
[17] See Bresnahan (1982) and Lau (1982).
[18] Empirical applications of the Bresnahan and Lau model to banking include e.g. Shaffer (1989, 1993, 2001), Berg & Kim (1994), Shaffer & DiSalvo (1994), Suominen (1994), Angelini & Cetorelli (2000), Coccorese (2005), Tsutsui & Uchida (2005).

4. The Methodology

This Chapter is divided into two sections. Section 4.1 outlines the empirical methodology used. Subsequently, Section 4.2 discusses some theoretical and methodological aspects of the methodology employed.

4.1 The empirical model

The literature on the measurement of market power in banking comprises structural as well as non-structural approaches. The structural approach to measure market power embraces the two competing hypotheses referred to as the Structural Conduct Performance paradigm, and the efficiency hypothesis. They are similar insofar as they both infer market performance from the level of market concentration (which is directly observable), but differ in their interpretations of the observed structure-performance relationship.

A crucial deficiency of structural approaches is their lack of theoretical foundation in oligopoly models where firm interactions come into play. Thus structural approaches rely on the strong assumption of Cournot conduct and hence are unable to reflect alternative oligopoly models. This deficiency triggered the development of various non-structural approaches,[19] namely the conjectural-variation model (Iwata, 1974), the Bresnahan-Lau model[20] and the Panzar-Rosse model[21]. In contrast to structural approaches, they base their performance inference on estimated conduct, not structure characteristics. An appealing feature of these models is the great amount of flexibility that they offer to the econometrician: Either we may proceed by considering conduct as a free continuous-valued parameter able to reflect any kind of oligopoly model on the continuum from perfect competition to monopoly, and subsequently apply a nested test to distinguish the consistent market hypothesis. Alternatively, we may fix the conduct parameter to certain values consistent with different game-theoretic models, and then distinguish the consistent oligopoly theory through a non-nested test method.

The paper relies on a non-structural approach, thus estimating conduct (market power) directly, as a free parameter. I employ a variant of the conjectural variation

---

[19] In this essay, the word "structural" refers interchangeably to (i) as a description of an approach which base its analysis on the market structure, and (ii) as a description of the model framework. While it is crucial not to confuse these interpretations, it should be clear from the context which interpretation is appropriate.

[20] See Bresnahan (1982) and Lau (1982).

[21] See Panzar & Rosse (1987).

(CV) approach, where conduct as well as cost is estimated efficiently.[22] Firms are assumed to maximize profits by setting equilibrium prices and quantities, subject to cost considerations and the degree of competition in the market. The degree of competition in turn depends on market demand characteristics and on firm behaviour (conduct). By assuming profit maximization, firm supply relations are derived, which together with a market demand function identify the conduct parameter. In similarity to the Bresnahan-Lau model, shifts in the exogenous market demand variables will trace out the supply relation consistent with a distinct level of market power.

Assume that the industry consists of *N* banks. Let q_i denote the amount of services produced by bank *i,* and $\sum_i q_i \equiv Q$ the quantity produced by the industry at market price P. Let the inverse market demand function be $P(Q,\mathbf{z})$, where $\mathbf{z}$ denotes a vector of exogenous variables affecting demand.

The cost function $C_i(\cdot)$ depends on the chosen scale of operation q_i and on the exogenously given prices of variable inputs $\boldsymbol{\omega}_i$.

Each bank carries out the following profit maximization program:

$$\underset{q_i}{Max}\ \pi_i = q_i P(Q,\mathbf{z}) - C_i(q_i,\boldsymbol{\omega}_i) \qquad [4.1]$$

The corresponding first-order condition is:

$$P = MC_i(q_i,\boldsymbol{\omega}_i) - q_i \frac{\partial P}{\partial Q}\frac{\partial Q}{\partial q_i} \qquad [4.2]$$

where $MC_i(q_i,\boldsymbol{\omega}_i)$ is the marginal cost of bank *i*. The second term on the right hand side measures the departure from marginal cost pricing. This term equals the product of the inverse of the market demand semi-elasticity to market price, i.e. $((\partial Q/\partial P)/Q)^{-1}$ and the conjectural elasticity, i.e. $\lambda_i = \frac{\partial Q}{\partial q_i}\frac{q_i}{Q}$.

The supply relation [4.2] nests all the standard oligopoly models (perfect competition; Cournot competition; and joint monopoly). If the market is perfectly competitive, a single bank anticipates an offsetting reaction from the other firms, leaving the total amount of services produced by the industry unchanged, i.e. $\partial Q/\partial q_i$=0. If firms were Cournot oligopolists, a single firm would not expect a

[22] See e.g. Appelbaum (1982).

reaction from the other firms to a change in its own output. Hence, the conjectural derivative, $\partial Q/\partial q_i$ equals one, implying $\lambda_i = q_i/Q$. In the joint monopoly (perfect collusion) case $\partial Q/\partial q_i$ again equals one, and since $q_i = Q$, it follows that $\lambda_i = 1$.

The market demand function faced by the competitors is stated as:

$$\ln Q = \alpha_0 + \alpha_1 \ln P + \alpha_2 \ln Z + \alpha_3 \ln Y + \varepsilon \qquad [4.3]$$

where ε is an error term and Y and Z are exogenous demand shifters. I follow common practice and include a variable that proxies for the level of general economic activity (*Y*), as well as a variable that proxies for the price of a substitute for banking services (*Z*), in the market demand equation.[23] Definitions of these variables appear in Table 5.1 below.

A necessary and sufficient condition for identification of λ in the simultaneous equation system estimated below is that the demand function must not be separable in at least one of the exogenous variables that are included in the demand equation but excluded from the marginal cost function (Lau, 1982). The identification condition is fulfilled in equation [4.3].[24]

A common choice of cost function in the analysis of banking markets is the transcendental logarithmic (translog) cost function. This is a flexible specification which avoids strong assumptions about the functional form. As a second order Taylor expansion in output and input levels it is able to approximate any twice differentiable function to the second degree.

$$\begin{aligned}\ln C_i &= c_0 + s_0 \ln q_i + \frac{s_1}{2}(\ln q_i)^2 + \sum_{j=1}^{3} c_j \ln \omega_{ji} + \ln q_i \sum_{j=1}^{3} s_{j+1} \ln \omega_{ji} \\ &\quad + \sum_{j=1}^{3}\sum_{k=1}^{3} c_{jk} \ln \omega_{ji} \ln \omega_{ki} + u_i\end{aligned} \qquad [4.4]$$

where u_i is an error term.

The implied marginal cost function is given by:

$$MC_i(q_i, \bar{\omega}_i) = \frac{\partial TC_i}{\partial q_i} = \frac{TC_i}{q_i}\left(s_0 + s_1 \ln q_i + \sum_{j=1}^{3} s_{j+1} \ln \omega_{ji} \right) \qquad [4.5]$$

It is not possible to predict the sign of the coefficients of the variables in translog cost function, but usually some restrictions are imposed in order to satisfy the

---

[23] See Bresnahan (1989) and Shaffer (1993).

[24] Consider, for instance, the case of *Z*. Since $\partial^2 Q/\partial P \partial Z = (\alpha_1 \alpha_2 Q)/(PZ) \neq 0$, the demand function is not separable in *Z*.

properties of a proper cost function[25]. Symmetry in the coefficients of produced goods is ruled out by the fact that only one (composite) product is considered. Symmetry in the coefficients of input prices would be necessary if we estimate different parameters for e.g. $\ln\omega_1 \ln\omega_2$ and $\ln\omega_2 \ln\omega_1$, rather than only one coefficient for each pair as in [4.4]. The properties of concavity and monotonicity do not constrain the coefficients on the terms involving $\ln\omega_{m,i}$ in the cost function. Linear homogeneity in input prices would imply $\sum_{j=1}^{3} c_j = 1$; $\sum_{j=1}^{3} s_{j+1} = \sum_{j=1}^{3}\sum_{k=1}^{3} c_{jk} = 0$.[26]

The supply relation [4.2] associated with [4.3] and [4.5] is:[27]

$$p_i = \frac{C_i}{q_i}\left(s_0 + s_1 \ln q_i + \sum_{m=1}^{3} s_{m+1} \ln\omega_{m,i}\right) - \lambda(P/\alpha_1) + \gamma_i \qquad [4.6]$$

where γ_i is an error term.

The parameter λ is identified in the system {[4.3];[4.6]}. By exploiting the cross-equations restrictions between equations [4.4] and [4.6], efficiency of the estimated parameters should improve (as the degrees of freedom increases).[28] Hence the simultaneous equation system to be estimated is comprised by equations [4.3], [4.4] and [4.6].

In view of the aggregate measure of output employed (cf. Section 5.2), it should be noted that the estimated behavioural parameter λ reflects the average conduct over the separate product markets, as well as over the years covered by the sample. Likewise, if banks enjoy varying degrees of market power, λ would reflect the behaviour of the average sample bank. Note also that the interpretation of λ as a measure of average conduct is valid regardless if the market was in equilibrium or disequilibrium during the period under study (Shaffer, 2001).

The panel data set also allows us to identify different conduct parameters for different groups of banks. For this purpose, the sample is partitioned according to ownership characteristics (commercial and savings banks) as well as branch network

[25] See Berger *et al*. (1987).
[26] I tested for linear homogeneity in input prices after performing the regressions. A chi-square test rejected linear homogeneity at the 1% level.
[27] I use the firm-level price as the dependent variable in [4.6], as suggested by Shaffer (1999). The advantage of using p_i instead of the aggregate market price *P* is that p_i reflects differences in the pricing structure (fee and commissions income vs. interest income) chosen by the different banks, and thus incorporates more information.
[28] See Bresnahan (1989), p 1040.

characteristics (unit banks and branch banks). A unit bank is a bank which has only one office, while a branch bank has more than one office.

A priori, it is difficult to unambiguously predict whether commercial or savings banks enjoy the highest price-marginal cost margins: On the one hand, cost-efficiency studies tend to provide evidence that savings banks are more non-interest cost inefficient than commercial banks, i.e. they use relatively more of inputs that are directly controllable by the management, such as labour and physical capital.[29] This suggests that commercial banks enjoy an advantage on the production side, implying superior performance, all else equal. In addition, commercial banks may have easier access to the capital market, enabling them to cut down on their finance cost.

On the other hand, if savings banks pursue additional objectives to profit maximization, markets where savings banks are present (i.e. local markets) are more likely to see high profits persist, because the competitive mechanisms of entry and exit are likely to be weak or inoperative (Goddard et al., 2004). This suggests that local markets are relatively more protected from competition than national or international banking markets. Recent studies provide empirical evidence that national and international markets are in fact more competitive than local markets.[30] According to the results in Bikker and Haaf (2002), this holds true for Sweden as well. However, it should be noted that the sample period is not up-to-date (1989-1998).

By estimating the simultaneous equation model separately for commercial banks and savings banks both average conduct and cost parameters are allowed to differ between the two groups.

The choice whether to be a unit bank or a branch bank is relevant in business activities where network size effects are important, such as retail banking. *A priori*, branch banks are expected to achieve superior performance. One argument for the superiority of branch banks given today is that they are able to diversify their asset portfolio. In addition, branch banks are able to reallocate capital from urban to rural areas at low cost, whereas unit banks typically have to raise all their capital and issue all of their loans locally (Seltzer, 2000). Furthermore, international research has

[29] See e.g. Hasan & Lozano-Vivas (2002).
[30] See DeBandt & Davis (2000) and Bikker & Haaf (2002).

shown that unit banks and branch banks face different cost structures and that branch banks tend to operate more efficiently.[31]

4.2 Methodological aspects

Although this methodology has been widely used in empirical applications, it nevertheless has come under question on theoretical as well as econometric grounds. One issue concerns the functional-form assumptions put on the model. The lesson from some previous studies which have imposed strong *a priori* functional-assumptions on e.g. demand is that it should be avoided, since otherwise inferences about market power or marginal cost may be incorrect.

Another source of criticism addresses the issue of the static nature of the conjectural variation approach. Opponents to the approach (e.g. Friedman, 1983) have argued that a dynamic interpretation of the conduct parameter is inconsistent with a model that is essentially static. However, the equilibrium of the static game may be regarded as the steady-state equilibrium of a corresponding dynamic game. Dockner (1992) proved that any CV equilibrium of a static game is equivalent to a steady state, sub-game perfect equilibrium of a dynamic game, and so concluded that a static CV analysis is justified for modelling dynamic interactions.

A third issue concerns the interpretation of the conduct parameter as a conjectural variation coefficient. Studies relying on the Bresnahan-Lau methodology assume that firms carry out their profit maximizations programs according to the conjectural variation (CV) model, although nothing in the empirical models employed actually restricts the interpretation of the estimated conduct parameter to be consistent with that of CV models (Bresnahan, 1989). Hence, the generated equilibrium may not necessarily be of the CV variety. On the other hand, Corts (1999) raises the fundamental objection that unless the observed equilibrium is of the static conjectural variety, the equilibrium *average level* of the price-marginal cost margin and the estimated *marginal* response in mark-ups to demand shocks will typically not be identical. In such cases, inferring the former from the latter (as we typically do) would be misleading. In particular, the estimated conduct parameter will underestimate the level of market power if the equilibrium is generated by certain dynamic oligopoly games, where demand shocks are not fully permanent.

[31] See Shaffer (1997).

Having Corts' critique in mind, Genesove & Mullin (1998) performed a test of the accuracy with which the CV approach can provide market power estimates. By comparing direct measures of marginal cost and price-cost margins with their estimated counterparts, they were able to conclude that the CV approach is performing reasonably well. Although the CV approach did underestimate the conduct parameter, the difference was considered as minimal. The smallest deviation was obtained when conduct was estimated as a free parameter. Finally, the results were robust to the specified demand functional-form.

5. The data

This Chapter is divided into three sections. Section 5.1 discusses the choice of input variables. Section 5.2 discusses and motivates the choice of output measure, while Section 5.3 discusses the sample.

Details on the definitions of all variables are given in Table 5.1 while descriptive sample statistics are presented in Table 5.2

5.1 Input variables

A disputed issue in banking is whether deposits should be treated as an input or an output. In the context of banking market power, most studies have employed the intermediation model of a banking firm which treats deposits as an intermediate input, used in conjunction with other input factors in the production of loans as well as other interest-bearing assets.[32] Alternative approaches, such as the user-cost model or the value-added model recognize that some liability items may earn money for the bank before converted into asset items, and thus ought to be considered as part of the output mix, rather than as inputs.[33]

Here, the intermediation model is followed, where banks are assumed to produce assets and other services using three variable inputs: deposits, labor and physical capital. Embodied in the specification of λ is an assumption that the associated input prices are treated as exogenously given. This assumption is probably unquestionable as far as the markets for labour and capital are concerned, since banks compete for these inputs with many other firms in other industries. It may also be true for deposits, if the banks compete effectively with each other for funds, and/or if deposit-taking is

[32] See Klein (1971).

[33] See Freixas and Rochet (1997) for a discussion about alternative approaches.

under effective competitive pressure from alternative investment options (as was probably the case during the period under examination).

Table 5.1: Variables used in regression analysis

Variable	Definition and interpretation
Balance sheet items	
q_i	Total assets of firm i (TA_i) : used as a proxy for firm level output, reflecting the multi-product nature of modern banking
$Q \equiv \sum_{i=1}^{N} q_i$	Aggregate output (considered as exogenous from the banks' viewpoint)
Total deposits (TD_i)	Include: saving and customer deposits, interbank deposits and securities issued
BR_i	The average (during the year) number of branches chosen by firm i. Considered as predetermined when the output decision is made
EMP_i	The average (during the year) number of employees of firm i.
Items from the profit and loss account	
Total interest revenues (TIR_i)	Includes total interest earnings on loans, interbank assets and the bond portfolio
Total interest expenses (TIE_i)	Includes interest cost on all liabilities (savings, customer and interbank deposits)
Total non-interest revenues ($TNIR_i$)	Income from services
TC_i	Total costs, i.e. total operating costs including interest expenses and provision costs
Total labor costs (TLC_i)	Direct and indirect staff costs
Total capital costs (TCC_i)	Depreciation costs (operating costs excluding staff and interest/provision costs)
Composite variables	
p_i	$(TIR_i + TNIR_i)/q_i$: firm level price of output
P	$P \equiv \sum_{i=1}^{N} (p_i)/N$: Market price
ω_{1i}	TIE_i / TD_i : proxy for input price of funds
ω_{2i}	TLC_i / EMP_i : proxy for input price of labour
ω_{3i}	TCC_i / TA_i : proxy for input price of physical capital
Exogenous variables	
Y	Gross domestic product (GDP): proxy for general economic activity
Z	Interest rate on a 3-month Swedish Treasury bill (risk-free): proxy for the price of a substitute to the services offered by the bank

Notes: Subscript indicates firm-level variable. Time subscripts are omitted.

If, on the other hand, banks have monopsony power in the market for funds, this will be miss-attributed to the asset side, and so the estimated λ will overstate the degree of market power on the asset side (Shaffer, 1993). Thus conditioned on the presence of monopsony power, a finding of perfect competitive behaviour would in fact constitute an even stronger result against hypotheses involving market power on the asset side.

As the associated input costs are not publicly stated, they are calculated using *ex post* account items. Thus the input cost of funds ω_1 is measured as total interest expenses divided by total deposits (including interbank takings and securities issued); input cost of labor ω_2 is defined as the ratio of wage costs to the number of employees; and the input cost of capital ω_3 is calculated as the ratio of all operating costs (including depreciation) net of interest and wage costs, to the number of branches. Definitions of all variables appear in Table 5.1, while sample statistics of operational variables are shown in Table 5.2.

5.2 Output variables

The output price is defined as the sum of total interest revenues and revenue from services, divided by total assets. The inclusion of non-interest revenues into our price definition is intended to reflect the increased importance of income sources generated by non-intermediation activities (e.g. fee and commission income from off-balance sheet business). Such a broad price definition was employed by e.g. Angelini & Cetorelli (2000), and is valid under the assumption that the stock of total assets is a good proxy for the heterogeneous flow of services supplied by banks, which is unobservable in our dataset. If large banks (in terms of asset-backed activities) also are large providers of off-balance sheet services (as seems likely), the ignorance of non-interest income may generate au upward bias in estimated marginal cost, in turn distorting the estimated λ and hence market power inferences (DeYoung, 1994).

Since banks offer a mix of services, I use a broad definition of output, proxied by total assets.[34] Given that the cost function is homothetic (separable in output quantities and input prices) the aggregation of outputs into a scalar index such as total assets is consistent (Shaffer, 1993) and preferable to an analysis based on a disaggregated vector of outputs for mainly two reasons:[35] (1) a scalar index is able to incorporate the effects of jointness in production, and (2) it allows for the strong

[34] See e.g. Shaffer (1993); Shaffer (2001); Shaffer & DiSalvo (1994); and Angelini & Cetorelli (2000).
[35] See Shaffer and David (1991).

assumption that expansion paths must lie along a ray through the origin in output space to be relaxed.

Furthermore, the aggregate approach is able to incorporate the effects of strategic pricing across product markets, while a single-product approach would bias the estimate of λ if banks indeed act strategically across the different product markets.[36]

Table 5.2: Sample statistics
(Swedish banking market 1996-2002)

Variable	Mean	Std.Dev	Min	Max
Q	2232867	360033	1620762	2728603
q_i	24823	112568	10.41	945545
p_i	0.0710	0.0124	0.0376	0.1110
C_i	1276	5636	0.483	47021
TD_i	19544	88367	9.53	732461
BR_i	22.58	95.11	1.00	1077
EMP_i	418.1	1722	1.00	12930
$TNIR_i/(TIR_i+TNIR_i)$	0.134	0.0721	0.0002	0.476
ω_{1i}	0.0299	0.0116	0.0112	0.0965
ω_{2i}	0.427	0.0961	0.0000	1.04
ω_{3i}	0.0174	0.0077	0.0014	0.0704
$\ln Y$	14.50	0.0660	14.40	14.58
Z	0.0257	0.0088	0.0111	0.0352

Notes: All figures are calculated from data supplied by Statistics Sweden and the Swedish Riksbank. Apart from ratios, number of branches, number of employees and the interest rate, figures are in MSEK and converted into 1995 values using the Gross Domestic Product deflator.

5.3 The sample

The dataset consists of an unbalanced panel of 631 annual observations on the variables involved (cf. Table 5.1), covering on average 90 individual banks over the period 1996-2002. The sample of banks (shown in Table A5) represents more than 90% of the entire Swedish banking industry in terms of total assets, and is composed of domestic commercial and savings banks.[37][38] The commercial banks are essentially

[36] See Angelini & Cetorelli (2000). The main drawback of using a scalar index as an output measure is obviously the inability to identify different λ for the separate product markets. This would require a disaggregated multi-product analysis, an approach which however is very demanding in terms of data requirements. See e.g. Gelfand & Spiller (1987), Suominen (1994) and Berg & Kim (1996) for such applications.

[37] Membership banks are excluded. All foreign banks (branches and subsidiaries) are also excluded due to insufficient data.

[38] The savings banks may operate one or several offices. A bank that operates only one office is considered as a unit bank, otherwise it is considered as a branch bank (cf. Section 4.1).

made up of (1) banks with a nationwide network of branches[39], (2) young internet banks (operating only one office) and (3) converted savings banks. Hence the sample represents a heterogeneous cross-section of banks as reflected by, *inter alia*, the large variation of size among the sample banks. This is shown in Table 5.2, which reports descriptive sample statistics for the variables used in the empirical analysis.

6. Estimation and results

This chapter is divided into three sections. Section 6.1 discusses estimation issues. Section 6.2 presents and discusses the estimation results while Section 6.3 interprets the results.

6.1 Estimation issues

The system {[4.3];[4.4];[4.6]} is estimated simultaneously using both iterative nonlinear seemingly unrelated regression (SUR) and generalized method of moments (GMM).[40] The GMM estimator addresses the endogeneity of q_i, p_i and C_i, and is therefore theoretically preferable. However, a problem encountered when using an instrumental variable estimator such as GMM is the difficulty of selecting relevant instruments. Thus it might be the case that SUR, which does not require instrumental variables, produces better estimates.

GMM is a technique used for simultaneous estimation of the whole system of equations. Thus both dependent regressors and cross-equation correlation of the disturbances are taken into account, while SUR only accounts for cross-equation disturbance correlations in order to improve efficiency.

Alternatively, the simultaneous equation model could be estimated using three-stage least squares (3SLS) or full-information maximum likelihood (FIML). FIML assumes that the disturbances have a multivariate normal distribution, while 3SLS and GMM does not assume a specific distribution of the disturbances. With normally distributed disturbances, 3SLS has the same asymptotic distribution as FIML, which is efficient among all estimators. However, 3SLS is typically preferred to FIML because it produces consistent estimates even if disturbances are not normally distributed, and, in addition, is far simpler to compute. The main advantage of GMM

[39] These banks are Nordea, SEB, Svenska Handelsbanken and Swedbank. Note that the present study only considers the domestic business activities of these banks.
[40] See Hansen (1982).

Table 6.1: System estimation results for the whole sample

Regressor		SUR		GMM	
Variable	Parameter	Estimate	t-Stat.	Estimate	t-Stat.
Demand function (dependent var: $\ln Q$)					
Constant	α_0	-11.64	-11.67*	-8.473	-11.45*
$\ln P$	α_1	-0.3872	-13.79*	-0.6122	-33.87*
$\ln Z$	α_2	-0.05201	-2.200**	-0.05078	-8.321*
$\ln Y$	α_3	1.726	22.45*	1.467	26.70*
R^2		0.9463		0.9358	
Cost function (dependent var: $\ln TC_i$)					
Constant	c_0	7.365	12.33*	12.91	5.545*
$\ln \omega_{1i}$	c_1	2.188	10.81*	2.414	3.610*
$\ln \omega_{2i}$	c_2	2.288	10.91*	5.686	5.483*
$\ln \omega_{3i}$	c_3	1.151	9.900*	2.001	4.336*
$\ln \omega_{1i} \ln \omega_{2i}$	c_4	0.2155	9.557*	0.1853	2.622**
$\ln \omega_{1i} \ln \omega_{3i}$	c_5	0.09818	7.775*	0.1824	3.083*
$\ln \omega_{2i} \ln \omega_{3i}$	c_6	0.1639	3.829*	0.3307	2.538**
$(\ln \omega_{1i})^2$	c_7	-0.05265	-6.867*	-0.08680	-3.727*
$(\ln \omega_{2i})^2$	c_8	0.2684	10.29*	0.7531	4.299*
$(\ln \omega_{3i})^2$	c_9	0.5199	6.034*	0.07340	2.326**
R^2		0.9987		0.9977	
Supply relation (dependent var: p_i)					
TC_i / q_i	s_0	0.6134	17.91*	0.08177	0.745
$(TC_i / q_i) \ln q_i$	s_1	-0.00059	-0.501	0.00234	0.812
$(TC_i / q_i) \ln \omega_{1i}$	s_2	-0.04590	-8.045*	-0.1064	-6.815*
$(TC_i / q_i) \ln \omega_{2i}$	s_3	-0.04658	-8.701*	-0.1476	-6.197*
$(TC_i / q_i) \ln \omega_{3i}$	s_4	-0.03970	-10.57*	-0.08979	-7.039*
Conduct	λ	0.08912	12.85*	0.1507	26.65*
R^2		0.3964		0.4395	
McElroy's R^2		0.9963		-	
Instruments		-		Constant; lagged q_i, p_i and C_i (logs, levels); current and lagged ω_{1i}, ω_{2i}, ω_{3i} and cross-products (logs); current Y and Z (logs); current and lagged EMP_i and BR_i (logs); time dummies	

Notes: *,** and *** indicate significance at the 1, 5 and 10% level, respectively. The reported t-statistics are robust to heteroscedasticity. No. of observations = 631 (unbalanced panel) over 3SLS is that GMM can correct for heteroscedasticity and autocorrelation of unknown form.[41] While the issue of autocorrelation is difficult to handle in the context of a simultaneous equation model, heteroscedasticity are taken into account using White's heteroscedasticity-consistent covariance matrix estimator (White, 1980).

It should be noted that GMM estimation is based on asymptotic theory and therefore could be subject to finite-sample bias in the present case. This problem is aggravated if the instruments have low relevance (weakly correlated with the endogenous variables) which tends to be a common phenomenon in empirical applications (Stock *et al*., 2002). Unfortunately it is not straightforward how to detect and handle weak instruments in the GMM model.[42]
By comparing the GMM and SUR estimates we can at least get an indication of the sensitivity of the results to the endogeneity problem.

6.2 Estimation results and goodness of fit

Table 6.1 reports the SUR and GMM estimates for the system {[4.3];[4.4];[4.6]}. The goodness of fit of the whole system is very good, as reflected by the McElroy's R-square value shown in the next to bottom row. Overall, the goodness of fit of the individual equations is also satisfactory, as reflected by the R-square values reported.

As shown in Table 6.1, the coefficient estimates obtained from the two methods are largely similar in sign and significance. However, the magnitude of the coeffcient estimates varies somewhat. In particular, the SUR coefficient on $\ln P$, α_1, which represents the market price elasticity of demand, is substantially lower in magnitude than the corresponding GMM coefficient. The fact that the SUR model appears to underestimate the sensitivity of demand to changes in price indicates that price is negatively correlated with the error term in the demand equation so that endogeneity indeed is an issue. Because of this, I focus mainly on the GMM results, although I also report the SUR results for comparison purposes.

With respect to the demand equation, both estimators yield parameters that are all significant at least at the 5% level. The estimate of α_1 is negative - consistent with a downward-sloping market demand curve - and strongly significant - a result which

[41] See Greene (1997) for more details on these estimators.
[42] See Stock *et al*. (2002) and Wright (2003).

directly affects the possibility of estimating the conduct parameter λ with great precision. In terms of magnitude, the point estimates of α_1, -0.39 and -0.61 respectively, are comparable to the average loan and deposit market price elasticities obtained in Oxenstierna (2000), which equal –0.40 and –0.51, respectively. With reference to the Norwegian banking industry, Berg & Kim (1994) report an estimated market price elasticity of -0.31, for the period 1980-89. Subsequently, Berg & Kim (1996) report disaggregate loan market elasticities for Norwegian banks for the period 1988-91. Thus, with respect to retail loans, the reported price elasticity equals -0.90, while for corporate loans, it equals -0.86. With respect to the Finnish banking industry, Vesala (1995) reports a loan market price elasticity of equal to -0.55.

The coefficient on the interest rate variable *Z* is significantly negative (at the 5% level), as shown. As a proxy for the price of a substitute to banking services, a positive sign was expected.[43] However the magnitude of α_2 is much lower than that of α_1, as expected, indicating that demand for banking services is much more sensitive to changes in the market price of bank output than to changes in the price of the substitute.

With regard to the variable measuring general economic activity, *Y*, Table 6.1 reports a positive and strongly significant coefficient. The coefficient on this variable is expected to be positive since more prosperous general economic conditions boost the demand for various banking services.

The fit of the cost equation is very good, as shown. All coefficients are significant at the 5% level in both the GMM and SUR models. Although the fit of the supply equation is more modest, the conduct parameter λ is estimated with great precision. The sign and magnitude of this parameter (0.09 and 0.15, respectively) indicates that banks on average exercised a significant, albeit small, degree of market power.

The results from the sub-sample estimations appear in Tables A1-A4 in the appendix. In the sub-sample estimations the demand equation parameters are constrained to equal their full-sample values, in order to keep e.g. estimated market demand elasticity to price invariant across the groups. As mentioned in Section 4.1, cost parameters and interaction are estimated separately for each group.

[43] Shaffer (1993) reports a positive but insignificant coefficient on this variable, while Shaffer (1999) reports a significantly negative coefficient.

Table A1 reports the results for commercial banks. As mentioned in Section 5.3, this group is made up of banks that differ with regard to geographic business scope as well as business strategy pursued (online and branch banking vs. pure online banking). As shown, the goodness of fit of the system is still very good, as reflected by the McElroy's R-square value, despite the smaller sample size. The fit of the cost function is similar to the full-sample case, while the fit of the supply equation is somewhat worse. In the SUR model, several of the marginal cost parameters, as well as the conduct parameter, are insignificant. Both the SUR and GMM models yield estimates of the conduct parameter that are much lower in magnitude compared to the full-sample values. Thus, the parameter value generated by SUR is only 0.02 while the GMM value is 0.13. A striking difference between the SUR and GMM models is that the latter model tends to yield more precise coefficient estimates with respect to this group. However, as the sample of commercial banks is small, the GMM estimates should be interpreted with some caution in this particular case. (cf. Section 6.1).

Table A2 reports the results for savings banks. The goodness of fit of the system is again very good, as shown. Note that while the fit of the cost function is still very good, the fit of the supply equation has improved compared to the fit obtained in the full sample. Both models now yield coefficients that are estimated with good precision. In particular, the conduct parameter is estimated very precisely. In addition, the magnitude of the conduct parameter is consistently higher than that obtained in the full-sample. The SUR model yields a conduct parameter equal to 0.10 while the GMM value is 0.15.

Table A3 reports the results for unit banks, i.e. banks that operate only one office. The sample size is smaller than the sub-sample for savings banks while larger than the sub-sample for commercial banks. The overall fit of the system is again very good and most of the cost and marginal cost parameters are statistically significant in both models. In particular, the conduct parameter is estimated with great precision. The estimates of the conduct parameter are similar in magnitude to the full-sample values (0.09 and 0.15).

Finally, Table A4 reports the results for branch banks. The sample size is larger than the sub-sample for unit banks, constituting approximately 60% of the full-sample. The system fit is very good and almost all coefficients are statistically significant. The estimates of the conduct parameter are somewhat higher in magnitude compared to the full-sample values (0.095 and 0.16).

6.3 Interpretation and discussion of results

As mentioned above, I focus on the GMM results, due to the presence of the endogeneity problem. The full-sample GMM point estimate of λ is approximately equal to 0.15 (cf. Table 6.1). As this value is significantly different from zero (at the 1% level) the hypothesis of perfect competition is rejected (recall from Section 4.1 that $\lambda=0$ under perfect competition). Likewise, the hypothesis of perfect collusion (joint monopoly) is strongly rejected, as shown in Table 6.2.

Furthermore, the hypothesis of Cournot conduct is also rejected. To see this, note that the degree of market power possessed by the average bank under Cournot oligopoly is proportional to market concentration (see Cowling & Waterson, 1976). In the present case, this corresponds to a degree of market power, measured in terms of the percentage price-marginal cost margin (Lerner index), of 35%. The corresponding value of the conduct parameter is approximately 0.21.[44] As shown in Table 6.2, the point estimates of λ are significantly different from this value, implying that the hypothesis of Cournot conduct is rejected (at the 1% level). It can therefore be concluded that on average, conduct was more competitive than the Cournot model would imply.

Table 6.2: Tests against various market hypotheses

Equilibrium	Conduct	Lerner index	Wald stat.	Decision
Perfect competition	0	0 %	427.2*	Reject
Estimated performance	0.15	22%	0	Accept
Cournot-Nash	0.21	35%	91.57*	Reject
Perfect collusion	1	152%	> 10 000*	Reject

Notes: Calculations are based on the GMM estimates. * = significant at the 1% level.

Table 6.3 summarizes key performance measures for the full-sample as well as for the different groups. The calculated values for MC (marginal cost) and AC (average cost) refer to sample averages. Note that for all groups, AC exceeds MC somewhat, indicating some modest scale economies in the industry. According to the results, this is particularly true for commercial banks and branch banks.

[44] The degree of market power of the average bank in the industry under Cournot conduct is calculated as $L = HHI / \alpha_1$, where HHI is the Herfindahl-Hirschman index, defined as the sum of the squared market shares (in terms of total assets) of all banks in the market. The value of the conduct parameter consistent with L can be derived from the last term in equation [4.6], which converts to a Lerner index if divided by the average sample price $\overline{P}$. Solving for λ, we obtain $\lambda \approx 0.21$.

Table 6.3: Performance measures

Sample	P_i	MC	AC	λ	L (%)
Full-sample	0.07101	0.05529	0.05642	0.1507	22.1
Commercial	0.06510	0.05967	0.06265	0.1273	8.34
Savings	0.07183	0.05481	0.05556	0.1526	23.7
Unit	0.06867	0.05455	0.05535	0.1511	20.6
Branch	0.07249	0.05507	0.05710	0.1597	24.0

Note: Calculations are based on the GMM estimates.

The group-specific Lerner indexes (*L*) reported in the far right column are calculated as the difference between the group-specific average price and the group-specific MC, divided by the group-specific average price. With regard to the full-sample, a Lerner index of approximately 22% is obtained, regardless if the calculations are based on the GMM estimates or the SUR estimates.[45]

The obtained value of *L* is comparable in magnitude to the value reported in Maudos & Nagore (2005). In their study, covering the period 1995-1999, Lerner indexes and marginal cost are estimated using samples of banks from 58 countries. With regard to Sweden, a Lerner index equal to 20% is reported.[46] The marginal cost reported equals 5.5%, which is very similar to the full-sample value reported in Table 6.3.

As reported in Section 6.2, both the GMM model and the SUR model indicate that average conduct and hence the average degree of market power (*L*) differs among the groups. For instance, the GMM model indicates that average conduct among savings banks was approximately 20% less competitive than average conduct among commercial banks, as the estimated conduct parameter for savings banks is 20% higher than for commercial banks. By means of chi-square tests, it was assessed that this difference is significant at the 1% level. Likewise, it was assessed that the difference in average conduct between branch banks and unit banks, amounting to +5%, is significant at the 5% level.

For all groups, the hypothesis of perfect collusion (joint monopoly) is strongly rejected. Likewise, the hypothesis of Cournot conduct is also rejected at the 1% level.

[45] *L* may be estimated without identifying α_1 and λ separately (see Appelbaum (1982) and Angelini & Cetorelli (2000)). In the present case, this involves simultaneous estimation of the system formed by equations [4.4] and [4.6]. Using this method, a value of *L* approximately equal to 22% is again obtained. Moreover, the sub-sample estimates of *L* are similar to the ones reported in Table 6.3.

[46] According to the results, this figure is higher than for the EU on average (15%) but similar to the world average (20%).

Overall, banks' behaviour was more competitive than what the Cournot model would imply, although not perfectly competitive.[47]

7. Conclusions

The objective of the paper has been to empirically assess the degree of competition in the Swedish banking market, over the period 1996-2002. For this purpose, a structural simultaneous-equation model, in which a conduct parameter is embedded, was estimated. The results indicate that the average bank's behaviour was more competitive than what the Cournot model would imply, although far from perfectly competitive. Furthermore, the results suggest that competition among commercial banks was significantly more intense than competition among savings banks. Put differently, commercial banks were operating in a more competitive environment than savings banks. This is consistent with the hypothesis of contestability – commercial banks operate in business segments where entry barriers are comparatively low. The results also suggest that banks operating more than one office experienced significantly better performance, in terms of price-cost margins, than one-office banks (unit banks).

[47] These findings are largely supported by the SUR regression model. The only difference is that the hypothesis of perfect competition cannot be rejected for the group of commercial banks.

Bibliography

Angelini, P & Cetorelli, N. (2000), ”Bank competition and regulatory reform: The case of the Italian banking industry” Working Paper Series, Federal Reserve Bank of Chicago, WP 99-32

Appelbaum, E (1982), “The estimation of the degree of oligopoly”, *Journal of Econometrics*, vol. 19, pp. 287-299

Bain, J.S. (1951), “Relation of Profit Rate to Industrial Organisation: American Manufacturing, 1936-40”, *Quarterly Journal of Economics*, 65, 293-324.

Baumol, W.J., Panzar, J.C., and Willig, R.D. (1982), *Contestable Markets and the Theory of Industrial Structure*, Harcourt Bruce Jovanovich, San Diego.

Berg, S.A. & Kim, M. (1996),”Banks as Multioutput Oligopolies: An Empirical Evaluation of the Retail and Corporate Banking Markets”, *Journal of Money, Credit, and Banking*, vol. 30, pp. 135-53.

Berg, S.A. & Kim, M. (1994),”Oligopolistic interdependence and the structure of production in banking: an empirical evaluation”, *Journal of Money, Credit, and Banking*, vol. 26, pp. 309-322.

Berger, A.N., Hanweck, G.A., and Humphrey, D.B. (1987), “Competitive Viability in Banking: Scale, Scope and Product Mix Economies”, *Journal of Monetary Economics*, vol. 16, pp. 501-20.

Berger, A.N. & Hannan, T.H. (1989), ”The Price-Concentration Relationship in Banking”, *Review of Economic and Statistics*, pp. 291-99.

Berger, A.N. & Hannan, T.H. (1991), “The Rigidity of Prices: Evidence from the Banking Industry”, *American Economic Review*, vol. 81, pp. 938-945.

Bikker, J.A. & Groenevald, H.J. (2000), “Competition and Concentration in the EU Banking Industry”, *Kredit und Kapital*, vol. 33, pp.62-98.

Bikker, J.A. & Haaf, K. (2002), ”Competition, Concentration and their Relationship: An Empirical Analysis of the Banking Industry”, *Journal of Banking and Finance*, vol. 26, pp. 2191-2214.

Bresnahan, T.F. (1982), “The oligopoly solution concept is identified”, *Economic Letters*, vol.10, pp. 87-92.

Bresnahan, T.F. (1989), “Empirical studies of industries with market power”, in Schmalensee, R. and Willig, R.D. (eds.), *Handbook of Industrial Organisation*, vol. II, North-Holland, Amsterdam.

Brozen, Y. (1982), *Concentration, Mergers, and the Public Policy,* Macmillan, New York.

Coccorese, P. (2005), "Competition in markets with dominant firms: A note on the evidence from the Italian banking industry", *Journal of Banking & Finance,* vol. 29, pp 1083-1093.

Coccorese, P. (1998), "Assessing the Competitive Conditions in the Italian Banking System: some Empirical Evidence", *BNL Quarterly Review*, vol. 51, pp. 173-191.

Corts, K. S. (1999), "Conduct Parameters and the Measurement of Market Power", *Journal of Econometrics*, vol. 88, pp. 227-250.

Cowling, K. & Waterson, M. (1976), "Price-Cost Margins and Market Structure", Economica, vol. 43, pp. 267-274.

DeBandt, O. & Davis, E.P. (2000), "Competition, contestability and market structure in European banking sectors on the eve of EMU", *Journal of Banking and Finance*, vol.24, pp. 1045-1066.

Demsetz, H. (1973), "Industry Structure, Market Rivalry, and Public Policy", *Journal of Law and Economics*, vol. 16, pp. 1-9.

DeYoung, R. (1994), "Fee-based Services and Cost Efficiency in Commercial Banks", Proceedings from a Conference on Bank Structure and Competition, *Federal Reserve Bank of Chicago*, pp. 501-519.

Dockner, E.J. (1992), "A Dynamic Theory of Conjectural Variations", *The Journal of Industrial Economics*, vol. 40, pp, 377-395.

ECON Report (2007), "Konkurrensen på Bankmarknaden", http://www.econ.se.

Evanoff, D. & Fortier, D. (1988), "Reevaluation of the Structure-Conduct-Performance Paradigm in Banking", *Journal of Financial Services Research*, vol. 1, pp. 249-60.

Freixas, X. & Rochet, J.C. (1997), *Microeconomics of Banking*, London: The MIT Press.

Friedman, J.W. (1971), "A Non-cooperative Equilibrium for Supergames", The Review of Economic Studies, vol. 38, pp. 1-12.

Friedman, J.W. (1983), *Oligopoly Theory*, Cambridge University Press, Cambridge.

Gelfand, M. & Spiller, P. (1987), "Entry Barriers and Multiproduct Oligopolies", *International Journal of Industrial Organization*, vol. 5, pp. 101-113.

Genesove, D. & Mullin, W.P. (1998), ”Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914”, *Rand Journal of Economics*, vol. 29, pp. 355-377.

Goddard, J.A. & Molyneux, P. & Wilson, J.O.S. (2004), “Dynamics of Growth and Profitability in Banking”, *Journal of Money, Credit, and Banking*, vol. 36, pp. 1070-1090.

Greene, W.H. (1997), “Econometric Analysis”, Prentice Hall.

Hannan, T.H. (1997), ”Market Share Inequality, the Number of Competitors, and the HHI“, *Review of Industrial Organization*, vol.12, pp. 23-35.

Hansen, L.P. (1982), “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica*, vol. 50, pp. 1029-1054.

Hasan, I. & Lozano-Vivas, A. (2002), “Organizational Form and Expense Preference: Spanish Experience”, Bulletin of Economic Research, vol. 54, pp. 135-150.

Iwata, G. (1974),”Measurement of Conjectural Variations in Oligopoly”, *Econometrica*, vol. 42, pp. 947-966.

Jackson, W.E.III (1992), “The Price-Concentration Relationship in Banking: A Comment”, *Review of Economic and Statistics*, vol. 74, pp. 373-376.

Jackson, W.E.III (1997), “Market Structure and the Speed of Adjustment: Evidence of Non-Monotonicity”, *Review of Industrial Organization*, vol. 12, pp. 37-57.

Klein, M. (1971), “A Theory of Banking Firm”, *Journal of Money, Credit, and Banking*, vol. 7, pp. 205-218.

Lau, L. (1982), “On Identifying the Degree of Competitiveness from Industry Price and Output Data”, *Economic Letters*, vol. 10, pp. 93-99.

Lybeck, J., *Den Svenska Finansiella Sektorns Utveckling under Modern Tid,* SOU, 2000:11, appendix 3.

Maudos, J. & Nagore, A. (2005), ”Explaining Market Power Differences in Banking: A Cross-country Study”, University of Valencia, Working Paper

Molyneux, P., Lloyd-Williams, D.M., and Thornton, J. (1994), “Competitive Conditions in European Banking”, *Journal of Banking and Finance*, vol. 18, pp. 445-59.

Neumark, D. & Sharpe, S. (1992), “Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposits”, *Quarterly Journal of Economics*, pp. 656-680.

Oxenstierna, G.C. (2000), ”Testing for Market Power in the Swedish Banking Oligopoly”, Working Paper 2/2000, Södertörns högskola.

Panzar, J.C. & Rosse, J.N. (1987), ”Testing for Monopoly Equilibrium”, *Journal of Industrial Economics*, vol. 25, pp. 443-456.

Rhoades, S.A. (1995), “Market Share Inequality, the HHI, and Other Measures of the Firm Composition of a Market, “*Review of Industrial Organization*”, vol. 10, pp. 657-74.

Seltzer, A.J. (2000), “A Comparison between Unit and Branch Banking: Australian Evidence o Portfolio Diversification and Branch Specialization, 1860-1930”, Working paper at Department of Economics, Royal Holloway College.

Shaffer, S. (1983), “Non-Structural Measures of Competition: Toward a Synthesis of Alternatives”, *Economic Letters*, vol. 12, pp. 349-53.

Shaffer, S. (1989), “Competition in the U.S. Banking Industry”, *Economic Letters*, vol. 29, pp. 321-23.

Shaffer, S. (1993), “A Test of Competition in Canadian Banking”, *Journal of Money, Credit, and Banking*, vol. 25, pp. 49-61.

Shaffer, S. (1997), “Network Diseconomies and Optimal Structure”, University of Wyoming, Working Paper.

Shaffer, S. (1999), “The Competitive Impact of Disclosure Requirements in the Credit Card Industry”, *Journal of Regulatory Economics*, vol. 15, pp. 183-98.

Shaffer, S. (2001), “Banking Conduct Before the European Single Banking License: A Cross-Country Comparison”, *North American Journal of Economics and Finance*, vol. 12, pp. 79-104.

Shaffer, S. & David, E. (1991), “Economies of Superscale in Commercial Banking” , *Applied Economics*, vol. 23, pp.283-93.

Shaffer, S. & DiSalvo, J. (1994), “Conduct in a Banking Duopoly”, *Journal of Banking and Finance*, vol. 18, pp. 1063-1082.

Smirlock, M. (1985), “Evidence on the (Non) Relationship between Concentration and Profitability in Banking”, *Journal of Money, Credit and Banking*, vol. 17, pp. 60-83.

Stock, J.H., Wright, J.H. & Yogo, M. (2002), “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments”, *Journal of Business and Economic Statistics*, vol. 20, pp. 518-29.

Suominen, M. (1994), “Measuring Competition in Banking: A Two Product Model”, *Scandinavian Journal of Economics*, vol. 96, pp. 95-110.

Tsutsui, Y. & Uchida, H. (2005), ”Has competition in the Japanese Banking Sector Improved?”, *Journal of Banking & Finance*, vol. 29, pp. 419-39.

Vesala, J. (1995), “Testing for competition in banking: Behavioral evidence from Finland”, Bank of Finland Studies, E:1.

White, H. (1980), “A Heteroscedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroscedasticity”, *Econometrica*, vol. 48, pp. 817-38.

Wright, J.H. (2003), “Detecting Lack of Identification in GMM”, *Economic Theory*, vol.19, pp. 322-330.

Appendix

Table A1: Results for commercial banks

Regressor		SUR		GMM	
Variable	Parameter	Estimate	t-Stat.	Estimate	t-Stat.
Demand function (dependent var: $\ln Q$)					
Constant	α_0	-11.64	[fixed]	-8.473	[fixed]
$\ln P$	α_1	-0.3872	[fixed]	-0.6122	[fixed]
$\ln Z$	α_2	-0.05201	[fixed]	-0.05078	[fixed]
$\ln Y$	α_3	1.726	[fixed]	1.467	[fixed]
Cost function (dependent var: $\ln TC_i$)					
Constant	c_0	7.079	4.926*	13.28	4.998*
$\ln \omega_{1i}$	c_1	2.264	3.323*	3.317	3.034*
$\ln \omega_{2i}$	c_2	3.002	5.506*	5.893	7.503*
$\ln \omega_{3i}$	c_3	1.433	7.409*	1.833	7.428*
$\ln \omega_{1i} \ln \omega_{2i}$	c_4	0.2579	2.789*	0.2964	2.594*
$\ln \omega_{1i} \ln \omega_{3i}$	c_5	0.05224	0.726	0.2786	2.690*
$\ln \omega_{2i} \ln \omega_{3i}$	c_6	0.3341	2.125**	0.5545	2.565**
$(\ln \omega_{1i})^2$	c_7	-0.004806	-0.304	-0.07449	-4.984*
$(\ln \omega_{2i})^2$	c_8	0.3973	5.019*	0.6071	4.283*
$(\ln \omega_{3i})^2$	c_9	0.1166	4.685*	0.03086	1.271
R^2		0.9992		0.9951	
Supply relation (dependent var: p_i)					
TC_i / q_i	s_0	0.8672	9.739*	0.2580	1.824***
$(TC_i / q_i) \ln q_i$	s_1	0.01962	1.623	-0.02142	-2.555**
$(TC_i / q_i) \ln \omega_{1i}$	s_2	0.01300	0.696	-0.1005	-2.971*
$(TC_i / q_i) \ln \omega_{2i}$	s_3	-0.009509	-0.363	-0.1369	-3.490*
$(TC_i / q_i) \ln \omega_{3i}$	s_4	0.01138	0.753	-0.1145	-8.272*
Conduct	λ	0.01921	1.337	0.1273	13.33*
R^2		0.2337		0.3318	
McElroy's R^2		0.9982		-	

Notes: *,** and *** indicate significance at the 1, 5 and 10% level, respectively. The reported t-statistics are robust to heteroscedasticity. No. of observations = 77 (unbalanced panel).

Table A2: Results for savings banks

Regressor		SUR		GMM	
Variable	**Parameter**	**Estimate**	**t-Stat.**	**Estimate**	**t-Stat.**
Demand function (dependent var: $\ln Q$)					
Constant	α_0	-11.64	[fixed]	-8.473	[fixed]
$\ln P$	α_1	-0.3872	[fixed]	-0.6122	[fixed]
$\ln Z$	α_2	-0.05201	[fixed]	-0.05078	[fixed]
$\ln Y$	α_3	1.726	[fixed]	1.467	[fixed]
Cost function (dependent var: $\ln TC_i$)					
Constant	c_0	6.453	3.989*	19.45	6.334*
$\ln \omega_{1i}$	c_1	0.6399	1.622	2.923	4.225*
$\ln \omega_{2i}$	c_2	1.658	6.380*	6.222	6.900*
$\ln \omega_{3i}$	c_3	1.713	4.081*	4.094	5.065*
$\ln \omega_{1i} \ln \omega_{2i}$	c_4	0.1261	4.500*	0.2191	3.653*
$\ln \omega_{1i} \ln \omega_{3i}$	c_5	0.09026	7.175*	0.2392	4.328*
$\ln \omega_{2i} \ln \omega_{3i}$	c_6	0.1489	3.522*	0.3631	3.260*
$(\ln \omega_{1i})^2$	c_7	-0.2510	-5.059*	-0.02127	-0.245
$(\ln \omega_{2i})^2$	c_8	0.1445	2.953*	0.7859	4.352*
$(\ln \omega_{3i})^2$	c_9	0.1779	5.882*	0.2516	4.517*
R^2		0.9978		0.9963	
Supply relation (dependent var: p_i)					
TC_i / q_i	s_0	0.3151	4.007*	-0.2961	-1.988**
$(TC_i / q_i) \ln q_i$	s_1	0.02166	6.817*	0.01982	4.218*
$(TC_i / q_i) \ln \omega_{1i}$	s_2	-0.03179	-3.569*	-0.1049	-6.233*
$(TC_i / q_i) \ln \omega_{2i}$	s_3	-0.03784	-4.088*	-0.1800	-6.999*
$(TC_i / q_i) \ln \omega_{3i}$	s_4	-0.09182	-8.618*	-0.1494	-7.193*
Conduct	λ	0.09774	49.30*	0.1526	50.78*
R^2		0.5646		0.5334	
McElroy's R^2		0.9939		-	

Notes: *,** and *** indicate significance at the 1, 5 and 10% level, respectively. The reported t-statistics are robust to heteroscedasticity. No. of observations = 554 (unbalanced panel).

Table A3: Results for unit banks

Regressor Variable	Regressor Parameter	SUR Estimate	SUR t-Stat.	GMM Estimate	GMM t-Stat.
Demand function (dependent var: $\ln Q$)					
Constant	α_0	-11.64	[fixed]	-8.473	[fixed]
$\ln P$	α_1	-0.3872	[fixed]	-0.6122	[fixed]
$\ln Z$	α_2	-0.05201	[fixed]	-0.05078	[fixed]
$\ln Y$	α_3	1.726	[fixed]	1.467	[fixed]
Cost function (dependent var: $\ln TC_i$)					
Constant	c_0	7.618	7.050*	9.376	3.301*
$\ln \omega_{1i}$	c_1	2.308	7.170*	1.616	2.214**
$\ln \omega_{2i}$	c_2	1.819	5.276*	4.317	3.601*
$\ln \omega_{3i}$	c_3	1.256	6.745*	1.462	2.610*
$\ln \omega_{1i} \ln \omega_{2i}$	c_4	0.2368	7.218*	0.09916	1.186
$\ln \omega_{1i} \ln \omega_{3i}$	c_5	0.05184	2.738*	0.1974	3.750*
$\ln \omega_{2i} \ln \omega_{3i}$	c_6	0.09120	1.412	0.2361	2.056**
$(\ln \omega_{1i})^2$	c_7	-0.01511	-1.005	-0.08719	-4.984*
$(\ln \omega_{2i})^2$	c_8	0.2741	9.850*	0.5171	2.659*
$(\ln \omega_{3i})^2$	c_9	0.04372	3.634*	0.05141	1.468
R^2		0.9988		0.9970	
Supply relation (dependent var: p_i)					
TC_i / q_i	s_0	0.6257	7.616*	0.1938	1.073
$(TC_i / q_i) \ln q_i$	s_1	0.005756	1.416	0.008873	1.424
$(TC_i / q_i) \ln \omega_{1i}$	s_2	-0.02390	-2.087**	-0.08988	-3.245*
$(TC_i / q_i) \ln \omega_{2i}$	s_3	0.001686	0.130	-0.1359	-3.642*
$(TC_i / q_i) \ln \omega_{3i}$	s_4	-0.05011	-7.711*	-0.06981	-4.077*
Conduct	λ	0.08907	22.22*	0.1511	28.64*
R^2		0.3599		0.3062	
McElroy's R^2		0.9969		-	

Notes: *,** and *** indicate significance at the 1, 5 and 10% level, respectively. The reported t-statistics are robust to heteroscedasticity. No. of observations = 245 (unbalanced panel).

Table A4: Results for branch banks

Regressor		SUR		GMM	
Variable	Parameter	Estimate	t-Stat.	Estimate	t-Stat.
Demand function (dependent var: $\ln Q$)					
Constant	α_0	-11.64	[fixed]	-8.473	[fixed]
$\ln P$	α_1	-0.3872	[fixed]	-0.6122	[fixed]
$\ln Z$	α_2	-0.05201	[fixed]	-0.05078	[fixed]
$\ln Y$	α_3	1.726	[fixed]	1.467	[fixed]
Cost function (dependent var: $\ln TC_i$)					
Constant	c_0	3.644	3.871*	12.40	6.058*
$\ln \omega_{1i}$	c_1	1.015	3.326*	1.771	2.589*
$\ln \omega_{2i}$	c_2	1.639	3.852*	6.793	8.026*
$\ln \omega_{3i}$	c_3	0.6932	3.170*	1.717	4.287*
$\ln \omega_{1i} \ln \omega_{2i}$	c_4	0.1366	4.339*	0.1214	1.767***
$\ln \omega_{1i} \ln \omega_{3i}$	c_5	0.3683	10.76*	0.6390	3.142*
$\ln \omega_{2i} \ln \omega_{3i}$	c_6	0.02164	0.269	0.3734	2.110**
$(\ln \omega_{1i})^2$	c_7	-0.1496	-9.111*	-0.1438	-4.536*
$(\ln \omega_{2i})^2$	c_8	0.09103	1.581	0.6772	4.946*
$(\ln \omega_{3i})^2$	c_9	0.05092	2.151**	0.02626	0.507
R^2		0.9985		0.9963	
Supply relation (dependent var: p_i)					
TC_i / q_i	s_0	0.7820	14.99*	-0.09608	-0.689
$(TC_i / q_i) \ln q_i$	s_1	-0.01543	-6.100*	-0.008033	-2.129**
$(TC_i / q_i) \ln \omega_{1i}$	s_2	-0.02770	-3.425*	-0.1059	-6.082*
$(TC_i / q_i) \ln \omega_{2i}$	s_3	-0.06709	-8.504*	-0.2282	-9.143*
$(TC_i / q_i) \ln \omega_{3i}$	s_4	-0.03711	-4.980*	-0.1342	-6.710*
Conduct	λ	0.09459	33.346*	0.1597	39.19*
R^2		0.4860		0.6064	
McElroy's R^2		0.9957		-	

Notes: *,** and *** indicate significance at the 1, 5 and 10% level, respectively. The reported t-statistics are robust to heteroscedasticity. No. of observations = 386 (unbalanced panel).

Table A5: Banks included in the study

Company name	Company name
Bergslagens Sparbank AB	Leksands Sparbank
Eskilstuna Rekarne Sparbank AB	Långasjö sockens Sparbank
Färs & Frosta Sparbank AB	Lönneberga Sparbank
FöreningsSparbanken Sjuhärad AB	Markaryds Sparbank
HSB Bank AB	Mjöbäcks Sparbank
IKANO Banken AB	Nordals Härads Sparbank
Kaupthing Bank Sverige AB	Norrbärke Sparbank
Länsförsäkringar Bank AB	Närs Sparbank
Nordea Bank Sverige AB	Orusts Sparbank
SkandiaBanken AB	Pitedalens Sparbank
Skandinaviska Enskilda Banken AB	Roslagens sparbank
Sparbanken Gripen AB	Röke Sockens Sparbank
Sparbanken i Lidköping AB	Sala Sparbank
Sparbanken Skaraborg AB	Sparbanken Nord
Stadshypotek Bank AB	Sidensjö Sparbank
Swedbank AB	Skatelövs och Västra Torsås Sparbank
Svenska Handelsbanken AB	Skurups Sparbank
Söderhamns Sparbank AB	Skånes Fagerhults Sparbank
Tjustbygdens Sparbank AB	Snapphanebygdens Sparbank
Varbergs Sparbank AB	Södra Dalarnas Sparbank
Vimmerby Sparbank AB	Sparbanken Finn
Ölands Bank AB	Sparbanken i Alingsås
Almundryds Sparbank	Sparbanken i Enköping
Alskogs Sparbank	Sparbanken i Ingelstorp
Attmars Sparbank	Sparbanken i Karlshamn
Bjursås Sparbank	Sparbanken Syd
Burs Pastorats Sparbank	Sparbanken Sörmland
Dalhems Sparbank	Sparbanken Tanum
Ekeby Sparbank	Sparbanken Tranemo
Eskelhems Sparbank	Sparbanken Västra Mälardalen
Falkenbergs Sparbank	Södra Hestra Sparbank
Fardhems Pastorats Sparbank	Sölvesborgs- Mjällby Sparbank
Farstorps Sparbank	Tidaholms Sparbank
Frenninge Sparbank	Tjörns Sparbank
Frykdalens sparbank	Tuna-Vena Sparbank
Garda-Lau Sparbank	Tyringe Sparbank
Glimåkra Sparbank	Ulricehamns Sparbank
Göteryds Sparbank	Vadstena Sparbank
Hishults Sparbank	Valdemarsviks Sparbank
Hudiksvalls Sparbank	Vallby Sparbank
Häradssparbanken i Mönsterås	Westra Wermlands Sparbank
Högsby Sparbank	Vinslövs Sparbank
Ivetofta sparbank i Bromölla	Virserums Sparbank
Järvsö Sparbank	Ydre Sparbank
Kinda sparbank	Ålems Sparbank
Kristianstads Sparbank	Åryds sparbank
Kräklingbo Sparbank	Åse och Viste härads Sparbank
Kyrkhults Sparbank	Åtvidabergs Sparbank
Laholms Sparbank	Älmeboda Sparbank
Lekebergs Sparbank	

Competition in Swedish Local Banking Markets

By

Pål Sjöberg[*]

Department of Economics and Statistics
Göteborg University

Abstract

While urban and metropolitan customers more and more use online banking as the principal delivery channel for accessing banking services, rural bank customers still rely on the physical network (branches) as the prime access channel. This means that high customer loyalty and entry barriers can be expected to prevail in rural banking. In light of this, the paper highlights rural banking conditions, and, more specifically, aims to evaluate the degree of competition in rural banking markets. For this purpose, a variation of the Bresnahan and Reiss (1991b) entry model is estimated using ordered probit and Poisson regression. According to the results, entry thresholds increase more than proportionately with each additional entry, suggesting that profit margins shrink as a result of new entry. The resulting pro-competitive effect is most pronounced in markets with a relatively few number of competitors, i.e. in markets accommodating fewer than the median number of four market players. Finally, the results suggest that a greater share of "multi-market banks" in a given market promotes local competition – a result which parallels a number of international studies.

JEL Classification: G21; L11; L13.

Keywords: Retail branch banking; Entry Barriers; Endogenous entry; Entry thresholds, Local banking markets; Ordered probit model, Event count model (Poisson)

[*] The author thanks Per-Olof Bjuggren, Dick Durevall, Lennart Flood, Lennart Hjalmarsson , Johan Lönnroth and Rune Stenbacka for useful comments. Financial support from the Swedish Competition Authority is gratefully acknowledged.

1. Introduction

According to a recent survey (Svenskt Kvalitetsindex, 2006) rural bank customers still rely on the physical network (branches) as the prime distribution channel for conducting retail banking services such as savings and lending, while for urban citizens online banking has become the prime channel. Amongst other things, this difference suggests that rural banking is still associated with high customer loyalty and persistent entry barriers, while in urban markets, by contrast, the extensive use of internet banking is expected to have reduced entry barriers and thus enabled increased competition. However, if competition indeed works improperly in Swedish rural banking, this could ultimately jeopardize the mission statement of the national program for rural development[48], i.e. an "ecologically, economically and sustainable development of the Swedish countryside", due to the role local banks play as the hub of the economy in rural parts. In light of this, I consider it to be of importance to undertake a study which examines competitive conditions in Swedish rural banking.

The study includes banks providing retail bank services to the countryside through a network of branches and which were in operation 1998-2002. The market is dominated by five large players, of which four are domestic and one is foreign. Their combined market share amounts to around 90 % in terms of total assets. These banks have in common a widespread network of branches covering an extensive geographic area, in some cases the whole country. The rest of the market is typically made up by comparatively very small banks that operate on a local scale. Nonetheless, these banks might enjoy a substantial or even dominant market share in their own local area. These banks are essentially made up of savings banks or former savings banks that have converted to joint stock companies.

The paper assumes that the market for retail branch banking is local in scope. Besides the growing body of international research supporting this view, the assumption can also be justified on the ground that the Swedish retail banking industry is partially comprised by local banks that often find themselves competing side-by-side with larger banks in terms of geographic scope (henceforth referred to as multi-market banks[49]). In addition, some of these multi-market banks explicitly

[48] The national program is part of a EU program for developing the countryside during the period 2007-2013.
[49] As far as I can figure out, the terms "multi-market bank" and "single-market bank" emanates from U.S. research within the same field. Since I find this terminology convenient, I adopt it. However, the

pursue a strategy tailored to fit local conditions, involving extensive local presence and decentralised decision-making.

For reasons explained later in the paper, local banking markets are proxied by local labour market areas. The median rural local labour market area has a population of 27 250, a number which corresponds fairly well to the median population of a rural U.S. county (24 000). U.S. rural counties have typically been used as spatial proxies for local banking markets, while in Europe the concept of local banking markets appears to be undeveloped.

A recent strand of the empirical industrial organisation literature uses game-theoretic entry models in order to estimate the degree of competition. The papers by Bresnahan & Reiss (1987, 1990, 1991b) and Berry (1992) are typically considered as the landmark work in this field. In this research context, the degree of competition is inferred from the impact additional entry has on the equilibrium price level (the toughness of price competition).

A simultaneous-move game theoretic model in the spirit of Bresnahan & Reiss (1991b) is employed, where competitors in the first stage decide on whether or not to enter a particular local market. Conditional on entry, they in the second stage participate in a price game (what matters are that they know the mode of game *ex ante*).

According to the results, "entry" of an additional firm always lowers per-firm margins (or raises the minimum per-firm market size[50] needed to cover fixed costs and thus be able to survive in long-rum equilibrium), suggesting the existence of rents that progressively diminishes with the entry of additional banks. Moreover, the results show that entry has the largest impact on per-firm margins in those markets that are most highly concentrated. In light of the results obtained we note that if the industry were already perfectly competitive, entry would not further reduce per-firm margins and, if incumbents were to collude, entry would again have no impact on per-firm margins. Per-firm margins that change as a result of entry indicate market equilibria characterized by an intermediate level of competition. It is a notable fact that the results obtained parallel those of Cetorelli (2002), who estimates entry thresholds for a cross-section of U.S. local rural banking markets, in the sense that competition

definition of these terms given later in the paper should not be confused with definitions appearing elsewhere.

[50] This mimimum level of per-firm market size is referred to as the per-firm entry threshold.

increases with every entry in the markets and that the competitive change is most pronounced in markets with two or three incumbents.

The paper concludes the following: (1) Swedish retail branch banking is a business that is indeed conducted locally; (2) profit margins in this industry appears to be quite high in the most concentrated local markets; (3) although margins are high in market with few competitors, they fall substantially with each additional entrant; (4) there is no empirical evidence of contestability; (5) there is no empirical support for collusive behaviour; (6) product differentiation appears to be limited and (7) a greater presence of large ("multi-market") banks seems to have a pro-competitive effect on profit-margins.

The rest of the paper is organized as follows. Chapter 2: section 2.1 reviews the recent literature on entry models and their applications to banking, while sections 2.2-2.3 are devoted to a discussion of the relevance of local banking with respect to the U.S. and Europe/Sweden, respectively. In Chapter 3 the methodology employed is thoroughly described, while chapter 4: section 4.1 describes and motivates the definition of local markets employed. With a proper market definition in place, section 4.2 describes and analyzes the market level data. Subsequently, in chapter 5, the results are presented and interpreted. Finally, chapter 6 concludes.

2. Background

2.1 Endogenous entry and banking – Review of the literature

Much of the recent literature on oligopolistic entry builds on the discrete-choice game theoretic entry models developed by Bresnahan & Reiss (1990, 1991a, 1991b). Within their original framework, homogeneous firms' profitability strategies are represented by discrete entry decisions that reflect underlying (1) market demand and cost characteristics; (2) the mode and intensity of post-entry competition; and (3) the simultaneous entry decisions of competitors. By assuming that firms enter only if they are able to make non-negative profits in equilibrium, the model provides a one-to-one-mapping from Nash equilibria to observed number of entrants (market structure). The role of the econometrician is to exploit information about latent firm profitability based on the endogenous market structure in order to infer the intensity of competition in the market.

Cetorelli (2002) adopts the methodology proposed by Bresnahan and Reiss (1991b) in order to assess the intensity of competition in U.S. local banking markets (rural counties). According to the results, entry into markets with the fewest number of competitors entails substantially larger pro-competitive effects (suggesting existing firms are making excessive rents) than in markets with five or more banks (suggesting that in rural banking, five competitors may be enough to achieve adequately competitive conditions).

The papers by Berry (1992) and Scott Morton (1999) extend the analysis by letting firm heterogeneity influence the entry decision. Building on these papers, Juan (2002) examines whether entry decisions are mainly influenced by local banking market conditions or firm characteristics. The study, which was carried out using data for the Spanish retail banking sector, essentially confirmed Sutton's "symmetry principle"[51] according to which firm heterogeneity (such as differences in size) does not have an impact on the entry decision. In contrast, the ambitious paper by Felici & Pagnini (2004) shows that both local market features and entrant characteristics are important. That is, based on a rather extensive analysis of the determinants of entry into a sample of Italian local banking markets, the authors were able to establish that larger banks have a higher probability of entering a new market.

The static entry models recently developed by Toivanen & Waterson (2001), Mazzeo (2002) and Seim (2004) allow for firm heterogeneity by endogenizing product type choice as well as the entry decision. For a cross-section of U.S. rural labour market areas, Cohen & Mazzeo (2004) examine competition among retail depository institutions using the Mazzeo (2002) framework. They consider a sequential-move Stackelberg game, where both the number and type of financial institutions (multi-market banks, single-market banks and thrift institutions) are included as arguments in the reduced-form profit function. In essence, their conclusions are twofold: 1) differentiation between the three different types is significant, and 2) the nature and extent of product differentiation depends on whether the market is more or less rural, in terms of the proximity to Metropolitan Statistical Areas.

2.2 Relevance of local banking: The case of the U.S.

Recent international research on the geographic scope of banking markets have found evidence that households and small businesses still tend to obtain several important

[51] See Sutton (1998).

financial services from nearby institutions. Moreover, costumers tend to cluster their purchase of financial services at a single financial institution [e.g. Kwast, Starr-McCluer & Wolken (1997); Amel & Starr-McCluer, (2002)]. Furthermore, several studies have investigated the relationship between deposit/loan interest rates and local market concentration. For example, Berger & Hannan (1989), Calem & Carlino (1991), and Hannan (1997), all find evidence of a negative relationship between local market HHI[52] and the deposit rate offered, suggesting that local market conditions determine banks' pricing policy.

On the other hand, Radecki (1998) shows that multi-market banks tend to offer uniform deposit interest rates across MSAs, suggesting that banking markets are not local in nature. Using similar survey data, Heitfield (1999) confirms these findings for multi-market banks, but also obtains significant differences in the deposit interest rates offered by single-market institutions across markets. Subsequently, Heitfield & Prager (2002) apply similar techniques to a much larger sample of banks covering a broader range of markets (i.e. urban as well as rural markets). Actually, for most types of deposit accounts, a significant negative relationship between local market concentration and price behaviour is obtained, reinforcing the view that local market conditions dictate banks' pricing behaviour.

The most recent research trend has been to focus on differences in pricing policy across different institution types. In this context, Hannan & Prager (2003) and Park & Pennacchi (2004) develop spatial models of bank pricing in order to analyze differences in pricing behaviour between multi-market and single-market banks. The findings of these two papers show that multi-market banks, due to their funding advantage vis-à-vis their single-market counterparts, tend to offer lower deposit interest rates *and* loan rates. Moreover, the larger the market share of multi-market institutions and the more concentrated the local market, the larger the impact of the funding advantage, and hence the lower the local market's equilibrium interest rates.

As long as the funding advantage is not offset by a loan operating cost disadvantage, a larger market share of multi-market banks promotes retail lending competition, while simultaneously harming competition in retail deposit business, leaving the overall effect ambiguous. However, in relatively concentrated local markets, the competitive effect on the loan side is likely to be more pronounced so

[52] Herfindahl-Hirshman concentration index.

that increased multi-market bank presence most likely reduces single-market bank profit and thus promotes local competition (Park & Pennacchi, 2004).

Hannan & Prager (2006) point out additional characteristics that potentially are relevant to the issue of how large bank competition affects local bank profitability: By virtue of their presence in many local markets, large banks may derive a benefit from geographic diversification, allowing them to offer lower loan rates for a given level of loan-specific risk, whereas local banks, on the other hand, have to collect their capital and issue their loans locally. In addition, by virtue of their size, large banks may offer a wider range of products and may as well offer products more efficiently than local banks. They conclude that, for rural banking markets, a greater presence of large banks is associated with a large, significant reduction in profitability for local banks. Furthermore, the magnitude of the reduction in profitability is larger the more concentrated the local market and the smaller the local banks are.

2.3 The relevance of local banking: The case of Europe and Sweden

A fundamental assumption underpinning the present framework is that the market for retail banking services that are provided through a network of branches indeed is local in scope. However, the local banking framework addresses two important issues: (1) the relevance of geographic proximity within retail banking. After the adoption of online banking, it may appear that geographic proximity has become less important. (2) No general clear-cut definition of a local banking market exists. Although there is empirical evidence related to the U.S. that banking markets such as retail banking are still local in scope, as outlined in section 2.2, the same issue very much remains an open question with respect to the euro area. This is understandable due to, on the one hand, the dearth of empirical evidence on this subject, and, on the other hand, to the fact that the European Commission so far has tended to assume that financial services such as retail banking are national in scope. This assumption can be explained by the absence of competition concerns in merger decisions relating to financial services, which has made a thorough analysis of the retail banking market unnecessary. However, the commission has left room for a regional definition in retail banking, and points out in its report[53] that at least some of the retail products (e.g. personal loans,

[53] See European commission paper (2006): *Interim Report II Current Accounts and Related Services* , available at:

small business banking) appear to be regional/local in scope, based on criteria such as the preference of banking customers for local suppliers, the significance of a dense branch network and the need for geographic proximity.

Recently, the Nordic competition authorities published the report "Competition in Nordic Retail Banking".[54] The report concludes that all Nordic retail markets are still dominated by a few large domestic banks with mainly loyal domestic clients. Although it is true that, on the one hand, markets have witnessed entry of new players (e.g. ICA bank and Danske Bank relating to the Swedish market), and, on the other hand, there has been increased investments in neighbouring countries, entry barriers are considered as persistently high, as well as concentration and profitability.

This result is even more striking in view of the increased reliance on internet banking. Internet banking reduces the need for a physical branch network and weakens customer-banking relationships. However, as pointed out in the report, retail banking belongs to a family of financial services that for households represent a substantial element of trust. Consumers may wish to remain with their well-known (local) providers of banking services despite better (but perhaps perceived as uncertain) deals being available. Such loyalty and consumer immobility constitute an entry barrier impeding effective (local) competition, and speaks to the existence of geographically delineated markets with rather few incumbents.

Unfortunately, the report is not explicit on the issue of geographic scope, with the exception of the Norwegian case. Thus, inquiries made by the Norwegian Competition Authority show that in competition for small and medium-sized enterprises, accessibility to the bank, local presence and a well established branch network are important features in the Norwegian retail banking market. Moreover, in relation to the DnB NOR merger case, it was established that the relevant market for most retail banking products were local or regional.

Turning to Sweden, the existent operation of the free-standing savings banks tacitly suggests the existence of local/regional banking markets, albeit yet to be defined. A recent survey by Svenskt Kvalitetsindex (2006) adds to the growing body of evidence that local banking is still highly relevant, despite the introduction of online banking. Thus, the survey measures the degree of client satisfaction as well as

---

http://ec.europa.eu/comm/competition/antitrust/others/sector_inquiries/financial_services/interim_report_2.pdf

[54] See Konkurrensverket (2006).

loyalty among bank customers in Sweden for the period 1997-2006. As it turns out, rural bank customers, who are relatively more prone to traditional branch banking[55], are more satisfied *and* more loyal than their urban counterparts. Put differently, those banks that pursue a business strategy involving branch banking and tailor their business strategy to fit local (rural) conditions, enjoy a higher degree of customer satisfaction and loyalty – a result which clearly adds to the conclusion that in Sweden local banking continues to play a vital role.

3. The methodology

3.1 Endogenous entry and ordered discrete-choice analysis

The aim of this section is to examine the *toughness of price competition*,[56] i.e. the relationship between the equilibrium price level and the number of competitors, in an industry that can be considered as offering homogenous services (Swedish retail branch banking), using a modified version of the entry model proposed by Bresnahan & Reiss (BR) in a sequence of papers (1987, 1990, 1991b). A two-stage game-theoretic model is employed, where banks in the first stage decide simultaneously on whether or not to enter a particular local market. Because the typical local market is fairly concentrated (median number of competitors is four) the market structure hypothesis postulates that markets are oligopolistic rather than competitive – implying that banks base their entry decisions not only on market demand and fixed costs of production, but on expectations about competitors' entry/operating decisions as well. Subsequently, conditional on entry, competitors play a price or quantity game (essentially, they know *a priori* the form of the game).

In essence, the toughness of price competition is inferred from the estimated relationship between the observed number of competitors and the minimum market size necessary to support the corresponding number of firms. Given an observed market structure, it is assumed that each incumbent is profitable, while the market cannot profitably support an additional entrant. In this way, the model provides a one-to-one mapping from Nash equilibria of the game to the observed number of firms.

The BR framework is essentially a static cross-sectional one, as it examines the existing market structure rather than actual entry. No distinction is made between

[55] Thus, according to the same survey, only 38% of free-standing savings banks' clients use internet as the prime channel. For other banks, internet has become the prime distribution channel.
[56] See Sutton (1991).

continuation of market operations and new entry. I follow this approach, using a pooled cross-sectional data set (1998-2002).

Let banks in a representative local banking market face a market demand function of the form[57]:

$$Q = d(p, \mathbf{x})S \qquad [3.1]$$

where $d(p, \mathbf{x})$ is the demand of a representative customer in the market; p is market price; $\mathbf{x}$ is a vector of exogenous market demand variables; and S is market size (population). It is assumed that S does not affect $d(p, \mathbf{x})$, so that a change in S correspond to a proportional change in total market demand, Q.

Given that n symmetric banks choose to enter the market, post-entry equilibrium profits for each bank will be given by:

$$\pi_n = (p_n - c_n)d(p_n, \mathbf{x}_n)\frac{S}{n} - F \equiv V_n \frac{S}{n} - F \qquad [3.2]$$

where c denotes variable costs; $(S/n) \equiv s$ is per-firm market size; V_n denotes per customer variable profits; and F is a fixed (sunk) entry cost.

An n[th] firm will choose to enter a market with *n-1* incumbents only if per-firm demand and hence variable profits is high enough to cover the fixed entry cost:

$$V_n s - F \geq 0 \qquad [3.3]$$

Since this expression is strictly increasing in per-firm market size, s, there exists a minimum value of s satisfying [3.3], for which n firms are just able to break even, corresponding to the zero-profit condition $V_n s - F = 0$. This per-firm market size level, denoted s_n, is known as the per-firm entry threshold. Solving for s_n:

$$s_n = \frac{F}{V_n} \qquad [3.4]$$

Thus s_n is increasing in F and decreasing in V_n. The intuition is clear: if entry causes the equilibrium price level and hence V_n to fall, banks need to compensate for this through a higher level of demand (per firm market size). Assuming a homogenous product industry, where subsequent entrants face the same cost structure, changes in s_n as n increases will be driven exclusively by changes in the price level and thus

[57] For notational convenience, market subscripts are omitted throughout this section.

variable profits, V_n. Given a downward sloping demand curve, the equilibrium price level p_n is a decreasing function of *n*, as long as incumbents do not perfectly collude.

If p_n were observable, the most straightforward way to assess the toughness of competition in the industry would by to estimate the relationship between p_n and *n*. However, since p_n is not observable at the disaggregated (local) market level considered here, the chosen framework utilizes that the (estimable) relationship between s_n and *n* may serve as a proxy for the (non-estimable) relationship between p_n and *n*. The estimated sequence of adjacent entry thresholds $s_1, s_2, ..., s_n$ reveals how additional entry affects profit margins (through p_n), and hence the degree of competition in equilibrium. Specifically, we infer changes in competition from adjacent entry threshold ratios s_{n+1}/s_n. If $s_{n+1}/s_n > 1$, entry of an n+1 firm has a pro-competitive impact, while if $s_{n+1}/s_n = 1$ competition does not change.

The estimated sequence of adjacent thresholds will trace out a path which is consistent with some mode of competition. It is instructive to analyze the implications of entry for three benchmark market structures: joint monopoly profit maximization; Cournot-Nash competition and perfect competition. If banks maximize joint monopoly profits, p_n will not be driven down as a result of additional entry, and hence s_n will remain unaffected as *n* increases. In a perfectly competitive industry, s_n will also be unaffected by additional entry since p_n is already at marginal cost level and cannot fall below this level post-entry.

In the case of Cournot behaviour, p_n must fall as a result of entry because incumbents do not change their output decisions as a result of entry. The prediction is that profit margins gradually decrease and s_n increase (at a decreasing rate), as *n* increases.

BR (1991b) estimates a profit function whose deterministic part is of the following form:

$$\bar{\pi}_{n,m} = V(n, \mathbf{x})S - F(\mathbf{w}) \qquad [3.5]$$

where the notations introduced above apply. $\mathbf{w}$ is a vector of per capita cost shifters. The effect of competition are accounted for by letting $V(\cdot)$ be a decreasing function of

n. $V(\cdot)$ and $F(\cdot)$ are specified as linear functions, which allow for a separate identification of the effect of entry on variable profits (*toughness of price competition*) versus fixed costs. The problem with the original BR specification is however that it is difficult to estimate; to separately identify variable profits and fixed costs parameters turn out to be difficult in practice.

For tractability purposes, the following reduced-form profit function specification is adopted:[58]

$$\pi_n = \overline{\pi}_n + \varepsilon \equiv \alpha \ln S + \mathbf{x}'\boldsymbol{\beta} - \mathbf{d}'_n \boldsymbol{\gamma}_n + \varepsilon \qquad [3.6]$$

where introduced notations still apply. $\overline{\pi}_n$ represents the deterministic part of profits while ε captures unobserved profits. ε is a market-level error term, assumed to follow a normal distribution, be additively separable from $\overline{\pi}_n$, independently distributed across markets, and identical for all banks within a given market. $\mathbf{d}_n$ is a vector of dummy variables indicating whether the number of banks in a given market equals n. α, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}_n$ are parameters to be estimated. The set of parameters $\boldsymbol{\gamma}_n$, which can be thought of as measuring the entry effect of the n^{th} bank on per-firm profits, are subsequently used to calculate entry thresholds. Market size S enters in log form[59], in order to ensure that the computed entry thresholds are non-negative.

In coherence with the entry model, potential entrants, at stage 1, simultaneously decide on whether to or not to enter. An nth entrant are assumed to enter if $E[\pi_n] \geq 0$, implying a certain outcome in terms of market players. In order to ensure that the generated outcome is consistent with Nash equilibrium, it is further assumed that:[60]

$$\pi_n \geq 0 \text{ and } \pi_{n+1} < 0 \qquad [3.7]$$

That is, if an outcome of n banks is observed, we infer that the preferred strategy by the n^{th} bank was to enter since it is able to make profits in equilibrium, while the preferred strategy of the $(n+1)^{\text{th}}$ was to stay out of the market.

---

[58] This specification is slightly different from the original BR model. Like the BR profit function, it can be interpreted as the log of demand (market-size) term multiplied by a variable profits term that depend on the number of market competitors (Mazzeo, 2002).
[59] Cf. e.g. Genesove (2004) and Cleeren *et al.*, (2006). Cleeren et al. (2006) remarks that the log form is consistent with a specification in which firms influence the ratio of variable profits to fixed costs, so that it becomes unnecessary to separately identify the effects of entry on variable profits and fixed costs, respectively.
[60] Cf. Bresnahan & Reiss (1991a).

Since π_n is a latent (unobserved) variable, we estimate [3.6] using an ordered probit model, where the dependent variable is the number of firms in a given market. The relationship between unobserved profits and the observed number of firms is given by:

$$\text{Number of firms} = n \text{ if } \gamma_n < \alpha \ln S + \mathbf{x}'\boldsymbol{\beta} + \varepsilon \leq \gamma_{n+1} \tag{3.8}$$

The probability of observing n banks in a given market m is:

$$P_{m,n} =$$

$$\begin{cases} P(\pi_2 < 0) = 1 - \Phi(\overline{\pi}_2) & \text{if } n \leq 1 \\ P(\pi_n \geq 0 \text{ and } \pi_{n+1} < 0) = \Phi(\overline{\pi}_n) - \Phi(\overline{\pi}_{n+1}) & \text{if } 0 < n < \hat{n} \\ P(\pi_{\hat{n}} > 0) = \Phi(\overline{\pi}_{\hat{n}}) & \text{if } \hat{n} \leq n \end{cases} \tag{3.9}$$

where $\Phi(\cdot)$ denotes the cumulative normal density function. Category $\hat{n}$ represents an aggregation of observations corresponding to n = 7 or more. This aggregation is made because of an insufficient number of markets with more than 7 banks; the ordered probit model requires a sufficient number of observations in each category.

The ordered probit model is estimated using maximum likelihood. As the ordered probit model only identify the parameters up to a scale factor, some normalization is required. I follow common practice and set the intercept term in $\boldsymbol{\beta}$ equal to zero. The log likelihood function to be maximized with respect to the elements of the vector $\{\alpha, \boldsymbol{\beta}\}$ along with the set of "cut points" $\{\gamma_2, \gamma_3, \ldots, \gamma_{\hat{n}}\}$ is then given by:

$$\begin{cases} \ln \ell = \sum_{m=1}^{M} \sum_{n=1}^{\hat{n}} d_{m,n} \ln(P_{m,n}) \\ \qquad\qquad \text{s.t.} \\ \gamma_1 = -\infty < \gamma_2 < \ldots < \gamma_{\hat{n}} < \gamma_{\hat{n}+1} = \infty \end{cases} \tag{3.10}$$

where the indicator variable $d_{m,n}$ equals one if n institutions are observed in market m and zero otherwise. The restriction is imposed in order to ensure that all probabilities given by [3.9] are positive. Assuming that the usual regularity conditions are fulfilled, the maximum likelihood estimators $\{\hat{\alpha}, \hat{\boldsymbol{\beta}}, \hat{\gamma}_n\}$ are consistent, asymptotically normal and efficient.[61]

[61] See e.g. Greene (1997).

3.2 Endogenous entry and event count analysis

As a complement to the ordered probit model, "entry thresholds" are also estimated using the log-linear Poisson model for count data.[62] The probability of observing n banks in market m is:

$$P(\text{No.of firms} = n) \quad = \frac{e^{-\lambda_m} \lambda_m{}^n}{n!}, \qquad \text{for } n = 0,1,2, \ldots \qquad [3.11]$$

where λ_m is the conditional mean parameter. The log-linear formulation implies that:

$$\lambda_m(\mathbf{x}_m, \boldsymbol{\delta}) = E(n \mid \mathbf{x}_m, \boldsymbol{\delta}) = \exp(\mathbf{x}'_m \boldsymbol{\delta}) \qquad [3.12]$$

where $\mathbf{x}_m$ is a vector including the same market characteristics as above and $\boldsymbol{\delta} \equiv \{\alpha_p \ \boldsymbol{\beta}_p\}$ is a vector of parameters to be estimated.

The parameters in [3.12] are estimated using maximum likelihood. Given independent observations, the log-likelihood function is (M denotes the number of markets):

$$\ln \ell(\boldsymbol{\delta}) = \sum_{m=1}^{M} \{ n\mathbf{x}'_m \boldsymbol{\delta} - \exp(\mathbf{x}'_m \boldsymbol{\delta}) - \ln n! \} \qquad [3.13]$$

The Hessian of $\ln \ell(\delta)$ is:

$$\frac{\partial^2 \ln \ell(\boldsymbol{\delta})}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} = -\sum_{m=1}^{M} \lambda_m \mathbf{x}_m \mathbf{x}'_m \qquad [3.14]$$

This expression is negative definite for all $\mathbf{x}$ and $\boldsymbol{\delta}$. That is, $\ln \ell(\boldsymbol{\delta})$ is globally concave, implying that convergence is guaranteed.

In order for statistical inference based on the maximum likelihood standard errors and t statistics to be valid, both the conditional mean λ_m and conditional variance must be correctly specified. For the Poisson maximum likelihood model this requires an assumption of equidispersion, that is, equality of λ_m and conditional variance. If this assumption fails to hold because the count data are overdispersed[63] (as is often the case), the Poisson maximum likelihood model still generates consistent estimates provided the conditional mean is correctly specified, but the standard errors will be severely downward biased (Cameron & Trivedi, 1998).[64] Test of overdispersion is performed using the likelihood ratio test provided in the Stata

[62] This idea follows Asplund & Sandin (1999).
[63] That is, the conditional variance exceeds the conditional mean.
[64] Cameron and Trivedi (1998) describe the restriction of equidispersion as qualitatively analogous to homoscedasticity in the linear model. However, a failure of the Poisson assumption of equidispersion may potentially have much larger effect on the estimated standard errors.

software package. Actually, the test does not reject the null hypothesis of equidispersion at any reasonable level of significance.

4. The data

This chapter is divided into two. Section 4.1 describes the nature of local banking in Sweden and provides arguments for the adopted local banking market definition. Given an accurate market definition, market level profitability is supposed to depend on factors that captures market demand and cost characteristics. These factors are proxied by observable variables that constitute a vector of exogenous market variables that enter the estimated reduced form profit function. The definition and description of the exogenous market variables is the subject of section 4.2.

4.1 Definition and analysis of Swedish local banking markets

The Swedish retail branch banking market is made up of, on the one hand banks that are purely local in geographic scope, and, on the other hand, banks that are active in a much wider geographic area, typically the whole country. Among the latter, some have chosen to pursue an explicit decentralized business strategy, tailored to fit conditions in local areas, while others' strategies are more or less centralized, involving uniform interest rates etc.

The retail branch banking market is dominated by four domestic large players: Svenska Handelsbanken (27%), SEB (24%), Nordea 16%) and Swedbank (15%).[65] Although these banks pursue different strategies, they have much in common. *Inter alia*, they all constitute financial conglomerates with a geographic scope extending beyond the Nordic countries. Furthermore, they all continue to rely on a widespread network of branches across the country. Branch banking is considered as an important complement to the ever-growing online banking user-base.

The fifth largest bank, Danske Bank i Sverige, is foreign and pursues a decentralized business strategy and operates through different province banks, such as Östgöta Enskilda Bank (Stockholm), Bohusbanken (Göteborg) , and Skånes Provinsbank.[66] The total market share of Danske Bank i Sverige is around 8% in total

[65] These figures are based on total assets, in 2004.
[66] The province banks are: Bohusbanken, Gävleborgs Provinsbank, Hallands Provinsbank, Närkes Provinsbank, Skaraborgs Provinsbank, Skånes Provinsbank, Smålandsbanken, Sundsvallsbanken, Sörmlands Provinsbank, Upplandsbanken, Värmlands Provinsbank, Västmanlands Provinsbank,

assets. The rest of the retail branch banking market is made up of saving banks or converted savings banks (joint-stock banks) which operate locally, or at most, regionally. There were 68 savings banks and 12 converted savings banks in operation at the end of 2006.

During the last decade, insurance and retail companies have founded niche banks such as Länsförsäkringar Bank, Skandiabanken, ICA-Banken and Ikano-Banken, which focus on the retail banking market. True enough, these internet-based banks have gained some limited market-share in certain segments at the expense of the four large players, but on an overall basis they are not considered as a serious threat since they do not offer a complete range of products. Of particular relevance for the present study is the limited adoption rate of internet-banking among rural customers (SKI, 2006), since predominance of branch banking more or less constitutes a necessary prerequisite for the present methodology to apply.

As clear from the foregoing market description, retail branch banks show a great deal of asymmetry with respect to geographic scope – some are purely local, while others are present in almost every local area across the country. This fact in itself clearly speaks to the existence of local banking markets, corresponding to independent geographical submarkets (albeit ambiguous and undefined) in Sweden, where branches of different banks offer competing products/services. In each independent submarket, different banks compete with each other through branches. Furthermore, the degree of substitution between banks belonging to the same submarket is expected to be quite high, reflecting fairly homogenous products and intra-market competition, while it should be zero or close to zero across markets, reflecting that submarkets are independent from the demand side.

Unfortunately, there exists no clear-cut universal definition of a local banking market, neither in Sweden nor elsewhere. With reference to the U.S., the Federal Reserve Banks broadly defines a local banking market as an economically integrated area that includes and surrounds a central city or a large town. In applied research, counties[67] have typically been considered as reasonable approximations of local banking markets. However, as remarked by Cohen and Mazzeo (2004), such political boundaries would be inappropriate if they do not represent meaningful economic

Älvsborgs Provinsbank and Östgöta Enskilda Bank. All these banks operate as independent units in their respective local markets.

[67] The median population of a U.S. county is around 24 000.

distinctions. They propose to use local labour market areas (LMAs), corresponding to independent geographical submarkets with respect to demand and supply of labour, as approximations of local banking markets. Based on commuting patterns between counties, U.S. LMAs are defined as integrated economic areas by the Bureau of Labour Statistics. Thus LMAs are recognized as functional territorial units, not administrative ones.

In the present paper, I adopt the Swedish correspondence of LMAs, i.e. the local labour markets (LLMs) developed by Statistics Sweden and ERU[68] as the basis for delineating local banking markets. The definition of LLMs[69] is based on municipalities as the smallest building blocks. Depending on the pattern of commuting streams, an LLM may correspond to a single municipality/ rural district, or, alternatively, involve a cluster of municipalities/rural districts. The grouping of municipalities into LLMs is shown in Table A3 in the appendix.

According to the division scheme based on the revision undertaken in 1998, (adopted here) the number of LLMs is 100. Historically, the LLM division scheme has been revised every fifth year, in order to reflect commuting streams in an up-to-date manner. The last decades have witnessed a fall in the number of LLMs, reflecting a process of extended commuting streams.

The three largest LLMs, i.e. the metropolitan areas of Stockholm, Gothenburg and Malmo are excluded from the study on the following grounds: (1) the study concerns rural conditions; (2) these markets are considerably larger than the median LLM and likely to contain distinct submarkets.

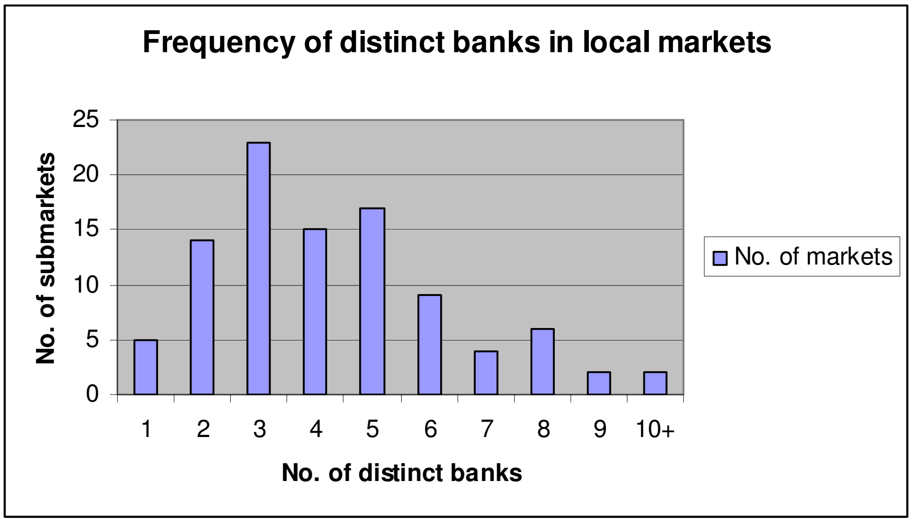
The median market population of the remaining 97 LLMs is 27 250, which corresponds fairly well to the median population of a U.S. county; 24 000.[70] The minimum population is 3 046 while the maximum population is 297 079. Only three of these markets have a population above 200 000 and a further 16 markets have a population above 100 000.

The number and identification of banks in a particular LLM is determined by aggregating the number of different banks that have a presence in the municipalities belonging to the LLM. As Figure 4.1 shows, a majority of the LLMs accommodate a

[68] Expertgruppen för forskning om regional utveckling.
[69] See the website: http://www.scb.se/templates/Standard____20125.asp for details about the criteria used to define the LLMs.
[70] As a further comparison, the Swedish counties (län) have a median population about ten times as large as that of the median LLM.

rather few number of banks. Thus, the median number of distinct banks is equal to 4, while more than 75% contain less than six banks. Apparently, this distribution is interesting from a competitive viewpoint, since *a priori* the fewer the market competitors, the easier it would be to successfully coordinate a cartel.

Figure 4.1

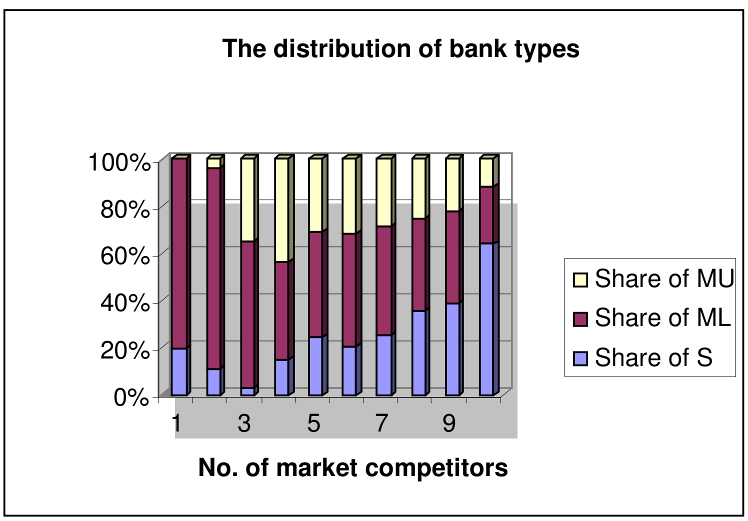


Source: Statistics Sweden, Branschregistret and own calculations

However, bank heterogeneity with respect to e.g. ownership type or geographic scope is likely to affect the assumed relationship between market structure and competition, and hence the likelihood of entry. For example, the savings (local) banks are distinct from their national commercial counterparts in that they pursue additional objectives to pure profit maximization. While true enough, this does not suggest that savings banks do not strive for high profits. After all, savings banks have no owners to turn to in order to raise new capital, so their only way to cope is to make enough profits. However, if indeed savings banks primarily pursue goals other than profit maximization, markets with a predominance of savings banks are more likely to see high (low) profits persist, because the equilibrating (competitive) mechanisms of entry and exit are likely to be weak or inoperative (Goddard *et al.*, 2004). Thus if anything, this suggests, *a priori*, that markets where only savings banks are present are less prone to competition than markets where commercial banks exert an influence.

As outlined in Chapter 2, there is a growing body of U.S. research investigating the impact of product differentiation (e.g. multi-market banks vs. single market banks) on market profitability and competition in retail banking. Whether the results obtained have relevance for markets in other countries is an empirical question left to be determined. At least it cannot be ruled out *a priori*. In order to account for potential differences in pricing behaviour between different types of banks, I classify banks into categories. Banks within a given category are assumed to be symmetric. A given bank is defined as a multi-market bank if it has a market presence in at least two distinct local markets, otherwise it is considered as a single-market bank (S). Among the group of multi-market banks, a further distinction is made according to whether the bank applies a uniform pricing across markets (MU), or pursue a localized pricing strategy (ML).[71] (The classification of individual banks into these categories is shown in Table A4 in appendix). Figure 4.2 below displays the distribution of different bank types in relation to the number of competitors. Apparently, multi-market banks are particularly prevalent, relatively speaking, in the most concentrated markets. If interpreted in accordance with Park & Pennacchi (2004), the scope for local banks to earn excess profits should be more limited than the market structure hypothesis would predict.

Figure 4.2



Source: Own calculations

[71] This distinction is not always trivial. In unclear cases, the distinction has been based on telephone interviews.

4.2 Exogenous market variables

Several local market characteristics may affect bank profitability and hence the likelihood of entry. Besides the level of population (in logs), market demand for banking services is expected to increase in per capita income, the employment rate, the number of farms and the number of local establishments. Per-capita income (*INC*) is measured as the pre-taxed yearly average labour income for the working-age population (20-64). The employment rate (*EMP*) is measured as the employed share of the working-age population. The number of local establishments *FIRMS* is included as a general measure of market prosperity. The number of farmings *FARMS* reflects the rural characteristics of many markets.

On the cost side, the rateable value for premises (*RENT*) is included. *RENT* is calculated as the ratio of total rateable values of all premises to the total area of premises, expressed in SEK/m^2. It is expected that entry should be relatively less likely in markets where *RENT* is high, all else equal.[72]

Data on these exogenous market variables were obtained from Statistics Sweden.[73] Table 4.1 reports descriptive sample statistics. It appears that the cross-sectional variation in the number of local establishments is particularly large.

Table 4.1: Sample statistics

Variable	Mean	Std.Dev	Min	Max
LPOP	10.3	1.15	8.02	12.6
INC	168	12.7	141	205
EMP	0.74	0.042	0.53	0.86
FARMS	1353	1253	79	5643
FIRMS	4193	4475	236	24404
RENT	2303	804	1115	5494

Notes: The sample consists of 97 markets, observed over the period 1998-2002.

Table 4.1 reports sample statistics of the exogenous market variables used. Data on these variables were collected for the period 1998-2002.

As reported above, recent international research (e.g. Park & Pennacchi, 2004; Hannan & Prager 2006) finds that greater presence of multi-market banks in rural,

[72] Other input costs, such as salary costs, are ignored because their cross-sectional variation is expected to be small.

[73] Data on *FIRMS* and *FARMS* were obtained from Statistics Sweden's Business Register, while data on *RENT* were obtained from Statistics Sweden's Register of Real Estate Assessments. The other data were taken from Statistics Sweden's website.

concentrated local markets promotes local competition by reducing the profitability of single-market banks. In essence, multi-market banks enjoy a funding advantage which enables them to exercise a competitive pressure upon their single-market counterparts.[74] In the spirit of this, I want to examine if market composition is an issue of relevance to competition also in Sweden. Thus I consider the following hypothesis: the greater the market-share of large banks pursuing a local presence strategy (MLs), the more competitive the market, *ceteris paribus*, and hence the lower the probability of entry. A binary variable *MULTI* is defined, which takes a value of one if a given market has a market-share of MLs exceeding the median value, and zero otherwise. *MULTI* enters [3.6] lagged by one year, consistent with a setting where potential entrants take market composition as exogenously given when they decide on whether or not to enter a market.

5. Results

The reduced form profit function [3.6] is estimated both with and without the inclusion of the composition indicator variable *MULTI*. Table 5.1 below reports the results from the ordered probit (OP) estimation without including this variable. As shown, most of the coefficients have their *a priori* expected signs, except for those of *INC* and *FARMS*. Thus higher population, employment rate and more local establishments increase bank profitability, in turn increasing the likelihood of entry, while an increase in the rateable value of premises lowers profits. However, only *LPOP* and *EMP* are significant at the 5% level.[75] The overall fit, in terms of pseudo R-square, is 0.32.

Table A1 in appendix reports ordered probit estimation results where *MULTI* is included. Interestingly, the coefficient of this variable is negative and significant, indicating that a greater market share of multi-market banks lowers profitability and hence the likelihood of entry, *ceteris paribus*. This result is clearly in accordance with the hypothesis stated above. As shown, the rest of the coefficients are only slightly modified. The results of the Poisson regression (PR) analysis are more or less similar, as shown in Table A2.

[74] Multi-market banks may enjoy additional advantages (cf. Hannan & Prager, 2006).

[75] One problem is that *LPOP*, *FIRMS* and *FARMS* are correlated with each other. The simple correlation between *LPOP* and *FIRMS* exceeds 0.8. In order to mitigate potential multicollinearity problems, I re-estimated the model with *FIRMS* and/or *FARMS* dropped. This change induced only a negligible effect on significant coefficients and overall goodness of fit.

Table 5.1: Results from ordered probit estimation

Parameter	Coefficient	Std. Err	z	P>|z|	[95% Conf. Interval]	
LPOP	1.43	0.129	11.11	0.000	1.18	1.68
INC	-6.33	4.74	-1.34	0.182	-15.6	2.96
EMP	5.47	1.45	3.78	0.000	2.64	8.31
FARMS	-0.0539	0.104	-0.52	0.605	-0.258	0.150
FIRMS	0.0284	0.0391	0.73	0.467	-0.0481	0.105
RENT	-0.139	0.108	-1.29	0.196	-0.351	0.0720
γ_2	14.27	1.338				
γ_3	15.63	1.367				
γ_4	17.19	1.427				
γ_5	17.95	1.439				
γ_6	18.89	1.452				
γ_{7+}	19.58	1.455				

Notes: All markets with 7 or more banks are aggregated into one category.
No. of obs = 485. LR chi2(6)= 584.91. Pseudo R^2=0.32. Log likelihood=-613.2.

The OP model generates estimated threshold parameters (cut points) $\hat{\gamma}_2, \hat{\gamma}_3, \ldots, \hat{\gamma}_{\hat{n}}$ which have no economic interpretation *per se* but can be used to calculate the more informative entry thresholds. Using [3.8] the minimum market size necessary to support *n* firms is given by:

$$S_n = \exp\left\{\frac{\hat{\gamma}_n - \bar{\mathbf{x}}'\hat{\boldsymbol{\beta}}}{\hat{\alpha}}\right\} \quad [5.1]$$

where regressors are set at their sample mean. Per-firm entry thresholds are calculated as $s_n = S_n / n$.

In the PR model, the necessary per-firm market size to support *n* firms is calculated using [3.11]. Evaluating covariates at their sample means, we obtain:

$$\lambda = n = \exp(\bar{\mathbf{x}}'_m \boldsymbol{\delta}) \quad [5.2]$$

from which S_n can be solved for.

The calculated entry thresholds appear in Table 5.2. Focusing first on the sequence of entry thresholds generated by the ordered probit model, it is clear that additional entry always promotes competition, since ratios always exceed one. Furthermore, the highest ratios correspond to the most concentrated markets (duopolies and triopolies), suggesting that the pro-competitive effect of entry is larger in these markets.

Table 5.2

Per-bank Entry threshold calculations						
Thresholds	**OP**	**OP per firm**	**OP per firm ratio**	**PR**	**PR per firm**	**PR per firm ratio**
1->2	3154	1577	-	5462	2731	-
2->3	8196	2732	1,73	14802	4934	1,81
3->4	24280	6070	2,22	30028	7507	1,52
4->5	41428	8286	1,37	51979	10396	1,38
5->6	79955	13326	1,61	81381	13564	1,30
6->7+	129330	18476	1,39	118884	16983	1,25

These results do not lend empirical support to the hypothesis of contestability.[76] If local markets were contestable, no distinct concentration-profit margins relationship would be obtained. Furthermore, the fact that each adjacent entry threshold is considerably larger than one is consistent with the assumption of a homogeneous product industry. However, in contradiction to the prediction of the Cournot model the relationship between entry threshold ratios and the number of firms does not describe a monotonically decreasing relationship. It is not straightforward to explain the non-monotonic pattern. For example, the fact that s_4 is considerably larger than s_3 is hard to reconcile with an explanation that the first three firms form a cartel, which a fourth entrant breaks up, since the arrival of a third entrant also entails considerable pro-competitive effects. At least we can conclude the following: (1) Profit margins must be quite high in the most concentrated markets, despite the heavy predominance of multi-market banks in these markets (cf. Fig. 4.2) and (2) there is no empirical support for the view that rural banks offer differentiated services.

As shown in the far right column of Table 5.2, the sequence of entry thresholds generated by the Poisson regression model describes a monotonically decreasing relationship between the number of firms and entry threshold ratios. In similarity with the "OP" sequence, threshold ratios are highest "in the beginning", subsequently falling gradually and rather slowly towards one. In essence, the conclusions drawn above are confirmed.

[76] A contestable market has low barriers to entry and exit. If the market is perfectly contestable, entry and exit are totally costless, i.e. there are no sunk entry costs. A high degree of contestability (potential competition) may render a market perfectly competitive regardless if the industry is highly concentrated or not. See Baumol *et al*. (1982).

6. Conclusions

While urban and metropolitan customers more and more use online banking as the principal delivery channel for accessing banking services, rural bank customers still rely on the physical network (branches) as the prime access channel. This means that high customer loyalty and entry barriers can be expected to prevail in rural banking.

In light of this, the purpose of the present paper has been to shed some light on the intensity (or lack of) of competition in Swedish retail branch (local) banking. For this purpose, a variation of the Bresnahan and Reiss (1991b) entry model was estimated using ordered probit and Poisson regression. According to the results, the following conclusions were drawn: (1) Swedish retail branch banking is a business that is indeed conducted locally; (2) profit margins in this industry appears to be quite high in the most concentrated local markets; (3) although margins are relatively high in market with few competitors, they fall substantially with each additional entrant; (4) there is no empirical evidence of contestability; (5) there is no empirical support for collusive behaviour; (6) product differentiation appears to be limited and (7) a greater presence of large ("multi-market") banks seems to have a pro-competitive effect on profit-margins.

Bibliography

Amel, Dean F. & Martha Starr-McCluer (2002) “Market Definition in Banking: Recent evidence”, *Antitrust Bulletin* vol.47, pp.63-88.

Asplund, M. & Sandin, R. (1999), ”The Number of Firms and Production Capacity in Relation to Market Size”, *The Journal of Industrial Economics*, vol. 47, pp. 69-85.

Baumol, W.J., Panzar, J.C., & Willig, R.D. (1982), “Contestable Markets and the Theory of Industrial Structure”, Harcourt Bruce Jovanovich, San Diego.

Berger, A.N. & Hannan, T.H. (1989), “The Price Concentration Relationship in Banking”, *Review of Economics and Statistics*, vol. 71, pp. 291-299.

Berry, S. (1992), “Estimation of a Model of Entry in the Airline Industry”, *Econometrica*, vol. 60, pp. 889-917.

Bresnahan, T. & Reiss, P. (1987), “Do Entry Conditions Vary Across Markets?”, *Brookings Papers in Economic Activities*, vol. 3, pp. 833-882.

Bresnahan, T. & Reiss, P. (1990), “Entry in Monopoly Markets”, *Review of Economic Studies*, vol. 57, pp. 531-553.

Bresnahan, T. & Reiss, P. (1991a), “Empirical Models of Discrete Games”, *Journal of Econometrics*, vol. 48, pp. 57-81.

Bresnahan, T. & Reiss, P. (1991b), “Entry and Competition in Concentrated Markets”, *Journal of Political Economy*, vol. 99, pp. 977-1009.

Calem, P.S. & Carlino, G.A. (1991), “The Concentration/Conduct Relationship in Bank Deposit Markets”, *Review of Economics and Statistics* vol. 73, pp. 268-276.

Cameron, A.C. & Trivedi, P.K. (1998), “Regression Analysis of Count Data”, Cambridge University Press.

Cetorelli, N. (2002), “Entry and Competition in Highly Concentrated Banking Markets”, *Economic Perspectives*, Q4, pp. 18-27.

Cleeren, K., Dekimpe, M.G., & Verboven, F. (2006), “Competition in Local-Service Sectors”, *International Journal of Research in Marketing*, vol. 23, pp. 357-367.

Cohen, A. & Mazzeo, M. (2004), “Market Structure and Competition among Retail Depository Institutions”, Finance and Economics Discussion Series 2004-4. Washington: Board of Governors of the Federal Reserve System, 2004.

European Commission Paper (2006), “Interim Report II Current Accounts and Related Services”.

Felici, R. & Pagnini, M. (2004), “Entry Decisions in Local Banking Markets: Does Potential Entrants’ Heterogeneity Matter?”, Mimeo, Rome, Bank of Italy.

Genesove, D. (2004), “”Why Are There so Few (and Fewer and Fewer) Two-Newspaper Towns?”, Mimeo, Hebrew University of Jerusalem, NBER and CEPR.

Goddard, J.A., Molyneux, P., & Wilson, J. (2004), ”Dynamics of Growth and Profitability in Banking”, *Journal of Money, Credit, and Banking*, vol. 36, pp. 1070-1090.

Greene, W.H. (1997), “Econometric Analysis”, Prentice Hall.

Hannan, T.H. (1997), “Market Share Inequality, the Number of Competitors, and the HHI”, *Review of Industrial Organisation*, vol.12, pp. 23-35.

Hannan, T.H. & Prager, R.A. (2003), “The Competitive Implications of Multi-market Bank Branching”, *Journal of Banking and Finance*, forthcoming.

Hannan, T.H. & Prager, R.A. (2006), “The Profitability of Small, Single-Market Banks in an Era of Multimarket Banking”, Finance and Economics Discussion Series working paper 2006-41, Board of Governors of the Federal Reserve System.

Heitfield, E. (1999), “What Do Interest Rate Data Say About the Geography of Retail Banking Markets?”, *Antitrust Bulletin* vol. 44, pp. 333-347.

Heitfield, E. & Prager, R.A. (2002), “The Geographic Scope of Retail Deposit Markets”, Finance and Economics Discussion Series working paper 2002-49, Board of Governors of the Federal Reserve System.

Juan, De R. (2002), “Entry in Independent Submarkets: An application to the Spanish Retail Banking Market”, *The Economic and Social Review*, vol. 33, pp. 109-118.

Konkurrensverket (2006), ”Competition in Nordic Retail Banking”, Report from the Nordic competition authorities 1/2006.

Kwast, M.L., Starr-McCluer, M. & Wolken, J.D. (1997), “Market Definition and the Analysis of Antitrust in Banking”, *Antitrust Bulletin,* vol. 42, pp. 973-995.

Mazzeo, M. (2002), “Product Choice and Oligopoly Market Structure”, *RAND Journal of Economics*, vol. 33, pp. 1-22.

Park, K. & Pennacchi, G. (2004), “Harming Depositors and Helping Borrowers: The Disparate Impact of Bank Consolidation”, University of Illinois Working Paper.

Radecki, L. (1998), "The Expanding Geographic Reach of Retail Banking Markets", *Federal Reserve Bank of New York Economic Policy Review,* pp. 15-33.

Scott Morton, F. (1999), "Entry Decisions in the Generic Pharmaceutical Industry", *Rand Journal of Economics*, vol. 30, pp. 421-440.

Seim, K. (2004), "An Empirical Model of Firm Entry with Endogenous Product-type choices", Stanford University Working Paper.

Sutton, J. (1991), "Sunk Costs and Market Structure", The MIT Press.

Sutton, J. (1998), "Technology and Market Structure. Theory and History", The MIT Press.

Svenskt Kvalitetsindex (2006), 2006-10-02.

Toivanen, O. & Waterson, M. (2001), "Market Structure and Entry: Where's the Beef?", *The Warwick Economics Research Paper Series*, No. 593, University of Warwick.

Appendix:

Table A1

Ordered Probit Estimation. Dep. Var. = No. of Market Competitors.

Parameter	Coefficient	Std. Err	z	P>|z|	[95% Conf. Interval]	
LPOP	1.36	0.131	10.37	0.000	1.10	1.62
INC	-5.65	4.77	-1.18	0.236	-15.0	3.70
EMP	4.95	1.46	3.40	0.001	2.10	7.81
FARMS	-0.0313	0.105	-0.30	0.767	-0.237	0.175
FIRMS	0.0228	0.0394	0.58	0.563	-0.0544	0.100
RENT	-0.108	0.109	-0.99	0.322	-0.105	0.321
MULTI	-0.635	0.112	-5.67	0.000	-0.854	-0.415
γ_2	12.93	1.376	(Ancillary parameters)			
γ_3	14.30	1.405				
γ_4	15.96	1.466				
γ_5	16.79	1.478				
γ_6	17.74	1.491				
γ_{7+}	18.45	1.493				

Notes: All markets with 7 or more banks are aggregated into one category.
No. of obs = 485. LR chi2(7)=617.27. Pseudo R^2=0.34. Log likelihood=-597.0.

Table A2

Poisson Maximum Likelihood estimation. Dep. Var. = No. of market players.

Parameter	Coefficient	Std. Err	z	P>|z|	[95% Conf. Interval]	
LPOP	0.371	0.0514	7.21	0.000	0.270	0.472
INC	-1.83	2.11	-0.87	0.387	-5.96	2.31
EMP	1.42	0.733	1.94	0.053	-0.0183	2.85
FARMS	0.00330	0.0358	0.09	0.927	-0.0668	0.0734
FIRMS	-0.00390	0.0118	-0.33	0.742	-0.0271	0.0193
RENT	-0.0573	0.0451	-1.27	0.204	-0.0312	0.146
MULTI	-0.185	0.0506	-3.66	0.000	-0.284	-0.0862
constant	-2.96	0.602	-4.91	0.000	-4.14	-1.78

Notes: No. of obs = 485. LR chi2(7)=370.16. Pseudo R^2=0.18. Log likelihood=-864.9.

Table A3: Local Banking Markets

Local Labour Markets	Municipalities included
Nyköping	Nyköping; Oxelösund
Katrineholm	Katrineholm; Vingåker
Eskilstuna	Eskilstuna; Flen
Linköping	Boxholm; Kinda; Linköping; Motala; Mjölby; Vadstena; Åtvidaberg; Ödeshög
Norrköping	Finspång; Norrköping; Söderköping; Valdemarsvik
Gislaved	Gislaved; Gnosjö; Hylte; Tranemo
Jönköping	Aneby; Habo; Jönköping; Mullsjö; Vaggeryd
Nässjö	Eksjö; Nässjö
Värnamo	Värnamo
Vetlanda	Sävsjö; Vetlanda
Tranås	Tranås; Ydre
Älmhult	Osby; Älmhult
Markaryd	Markaryd
Växjö	Alvesta; Lessebo; Tingsryd; Uppvidinge; Växjö
Ljungby	Ljungby
Hultsfred	Hultsfred
Emmaboda	Emmaboda
Kalmar	Borgholm; Kalmar; Mörbylånga; Nybro; Torsås
Oskarshamn	Högsby; Mönsterås; Oskarshamn
Västervik	Västervik
Vimmerby	Vimmerby
Gotland	Gotland
Olofström	Karlshamn; Olofström
Karlskrona	Karlskrona; Ronneby
Perstorp	Perstorp
Helsingborg	Bjuv; Båstad; Helsingborg; Höganäs; Klippan; Landskrona; Svalöv; Åstorp; Ängelholm; Örkelljunga
Kristianstad	Bromölla; Hässleholm; Kristianstad; Sölvesborg; Östra Göinge
Simrishamn	Simrishamn; Tomelilla
Halmstad	Halmstad; Laholm
Falkenberg	Falkenberg
Varberg	Varberg
Bengtsfors	Bengtsfors; Dals-Ed
Lysekil	Lysekil; Sotenäs
Strömstad	Strömstad; Tanum
Trollhättan	Färgelanda; Grästorp; Lilla Edet; Mellerud; Munkedal; Trollhättan; Uddevalla; Vänersborg
Borås	Borås; Herrljunga; Mark; Svenljunga; Ulricehamn
Lidköping	Essunga; Götene; Lidköping; Vara
Skövde	Falköping; Gullspång; Hjo; Karlsborg; Mariestad; Skara; Skövde; Tibro; Tidaholm; Töreboda
Torsby	Sunne; Torsby
Munkfors	Munkfors
Årjäng	Årjäng
Karlstad	Forshaga; Grums; Hammarö; Karlstad; Kil
Kristinehamn	Kristinehamn
Filipstad	Filipstad
Hagfors	Hagfors
Arvika	Arvika; Eda
Säffle	Säffle; Åmål
Laxå	Laxå
Hällefors	Hällefors
Örebro	Askersund; Hallsberg; Kumla; Lekeberg; Lindesberg; Nora; Örebro
Karlskoga	Degerfors; Karlskoga; Storfors
Västerås	Hallstahammar; Sala; Surahammar;Västerås
Fagersta	Fagersta; Norberg; Skinnskatteberg
Köping	Arboga; Kungsör; Köping
Vansbro	Vansbro
Malung	Malung
Älvdalen	Älvdalen
Mora	Mora; Orsa
Falun	Borlänge; Falun; Gagnef; Leksand; Rättvik; Säter
Avesta	Avesta; Hedemora
Ludvika	Ljusnarsberg; Ludvika; Smedjebacken
Hofors	Hofors
Ljusdal	Ljusdal
Gävle	Gävle; Ockelbo; Sandviken; Älvkarleby
Söderhamn	Söderhamn
Bollnäs	Bollnäs; Ovanåker
Hudiksvall	Hudiksvall; Nordanstig
Ånge	Ånge
Härnösand	Härnösand
Sundsvall	Sundsvall; Timrå
Kramfors	Kramfors
Sollefteå	Sollefteå
Örnsköldsvik	Örnsköldsvik
Strömsund	Strömsund
Åre	Åre
Härjedalen	Härjedalen
Östersund	Berg; Bräcke; Krokom; Ragunda; Östersund
Storuman	Storuman
Sorsele	Sorsele
Dorotea	Dorotea
Vilhelmina	Vilhelmina
Åsele	Åsele
Umeå	Bjurholm; Nordmaling; Robertsfors; Vindeln; Vännäs; Umeå
Lycksele	Lycksele; Malå
Skellefteå	Norsjö; Skellefteå
Arvidsjaur	Arvidsjaur
Arjeplog	Arjeplog
Jokkmokk	Jokkmokk
Överkalix	Överkalix
Kalix	Kalix
Övertorneå	Övertorneå
Pajala	Pajala
Gällivare	Gällivare
Älvsbyn	Älvsbyn
Luleå	Boden; Luleå; Piteå
Haparanda	Haparanda
Kiruna	Kiruna

Table A4: The banks included in the data set

Company name	Strategy
SKANDINAVISKA ENSKILDA BANKEN AB	MU
FÖRENINGSSPARB. SJUHÄRAD AB	S
NORDEA BANK SVERIGE AB (PUBL)	MU
LÄNSFÖRSÄKRINGAR BANK AB	ML
SVENSKA HANDELSBANKEN AB	ML
ESKILSTUNA REKARNE SPARBANK AB	S
FÖRENINGSSPARBANKEN ÖLAND AB	S
SÖDERHAMNS SPARBANK AB	S
BERGSLAGENS SPARBANK AB (PRIVAT)	MU
SPARBANKEN SKARABORG AB	ML
VARBERGS SPARBANK AB	ML
SPARBANKEN LIDKÖPING AB	S
VIMMERBY SPARBANK AB	S
TJUSTBYGDENS SPARBANK AB	S
SPARBANKEN GRIPEN AB	S
SWEDBANK AB	ML
DANSKE BANK I SVERIGE AB	ML
SPARBANKEN NORD	ML
ULRICEHAMNS SPARBANK	S
SALA SPARBANK	S
SPARBANKEN I KARLSHAMN	S
WESTRA WERMLANDS SPARBANK	ML
SPARBANKEN VÄSTRA MÄLARDALEN	S
FALKENBERGS SPARBANK	MU
SPARBANKEN SYD	S
KRISTIANSTADS SPARBANK	S
HUDIKSVALLS SPARBANK	S
LEKSANDS SPARBANK	S
SPARBANKEN TRANEMO	S
SÖDRA DALARNAS SPARBANK	S
NORDALS HÄRADS SPARBANK	MU
TIDAHOLMS SPARBANK	MU
FRYKSDALENS SPARBANK	S
VALDEMARSVIKS SPARBANK	MU
SÖLVESBORG MJÄLLBY SPARBANK	S
KINDA SPARBANK	S
ÅTVIDABERGS SPARBANK	S
HÖGSBY SPARBANK	S
ÅSE OCH VISTE HÄRADS SPARBANK	S
VADSTENA SPARBANK	S
MARKARYDS SPARBANK	S

Company name	Strategy
ALMUNDSRYDS SPARBANK	S
ALSKOGS SPARBANK	S
ATTMARS SPARBANK	S
BJURSÅS SPARBANK	S
BURS PASTORATS SPARBANK	S
DALHEMS SPARBANK	S
EKEBY SPARBANK	S
ESKELHEMS SPARBANK	S
FARSTORPS SPARBANK	S
GARDA-LAU SPARBANK	S
GLIMÅKRA SPARBANK	S
GÖTERYDS SPARBANK	S
HISHULTS SPARBANK	S
HÄRADSSPARBANKEN MÖNSTERÅS	S
IVETOFTA SPARBANK I BROMÖLLA	S
JÄRVSÖ SPARBANK	S
LEKEBERGS SPARBANK	S
KYRKHULTS SPARBANK	S
LÅNGASJÖ SOCKENS SPARBANK	S
LÖNNEBERGA SPARBANK	S
MJÖBÄCKS SPARBANK	S
NORRBÄRKE SPARBANK	S
NÄRS SPARBANK	S
RÖKE SOCKENS SPARBANK	S
SIDENSJÖ SPARBANK	S
SKATELÖVS OCH VÄSTRA TORSÅS SB	S
SKÅNES FAGERHULTS SPARBANK	S
SÖDRA HESTRA SPARBANK	S
TUNA-VENA SPARBANK	MU
TYRINGE SPARBANK	S
VALLBY SPARBANK	S
VINSLÖVS SPARBANK	S
VIRSERUMS SPARBANK	S
ÅLEMS SPARBANK	S
ÅRYDS SPARBANK	S
ÄLMEBODA SPARBANK	S
SNAPPHANEBYGDENS SPARBANK	MU
SPARBANKEN TANUM	S
LAHOLMS SPARBANK	S
SPARBANKEN SÖRMLAND	ML

Notes: S denotes a Single market bank; ML denotes a multi-market bank, pursuing a local strategy, while MU denotes a multi-market bank pursuing a uniform strategy.

A Dynamic Analysis of Firm Growth in Swedish Banking

By

Pål Sjöberg*

Department of Economics and Statistics
Göteborg University

Abstract

The objective of this paper is to contribute to the understanding of firm growth dynamics in the new banking environment, by testing the validity of Gibrat's Law of Proportionate Effect on Swedish data. The point of departure in the paper is the expectation that large banks should be able to more fully exploit scale and scope economies associated with technological innovations such as internet banking, than smaller banks, and therefore grow faster. However, such a predictable growth pattern would be inconsistent with Gibrat's law. Using a panel of 79 Swedish banks over the period 1995-2002, I find no empirical evidence that large banks grew faster, nor any significant evidence that firm sizes were mean-reverting. Hence the Law was not rejected. However, growth was not entirely random, as banks with a more diversified revenue mix experienced significantly higher growth rates than less diversified banks.

JEL classification: C23; G21; L11.

Keywords: Dynamic panel data model; Firm growth; Gibrat's law of proportionate effect; GMM estimation; Swedish banking; Technological change

* I am grateful to Dick Durevall, Mattias Ganslandt, Lennart Hjalmarsson, Johan Lönnroth and Rune Stenbacka for giving helpful comments and suggestions.

1. Introduction

Gibrat's Law of Proportionate Effect (LPE) states that if each firm's growth rate in a given period is independent of its size and growth in previous periods (or, equivalently, if logarithmic firm sizes are subject to sequences of purely random shocks), then the firm size distribution will become increasingly right-skewed, and dominated by a small number of large firms, eventually turning to the log-normal distribution.[77] It is well known that the empirical size distribution of firms in many industries, such as banking, resembles the lognormal distribution.

By contrast, other studies have proposed various factors that may explain market structure in banking such as economies of scale and scope, M&A activities, strategic behaviour of incumbent banks and technological innovations. Cost-reducing technological innovations can be expected to impact differently on large and small banks. Because large banks more easily undertake the large fixed costs associated with implementation of technological innovations, and thus should be able to fully exploit the associated economies of scale and scope, they are *a priori* expected to grow more rapidly than small banks. However, such a predictable growth pattern would contradict the LPE, which postulates that each bank's growth rate in any year is completely randomly determined.

The prime objective of the paper is to test the validity of the LPE in light of the implementation of new technology, including internet banking which was first launched in Sweden in 1995. Internet banking and other forms of electronic banking services have reduced transaction costs substantially and contributed to an average increase in productivity of 4.6% per year over the period under study (1995-2002).[78] However, since the gains in productivity were not evenly distributed across the population of banks, it is likely that growth opportunities were not either. If this conjecture is correct, the LPE should be rejected.

An additional objective is to examine if various bank-specific performance indicators have a systematic impact on growth rates.

The estimated model of firm growth contains a lagged dependent variable and individual bank-level effects. To obtain consistent (and asymptotically efficient) estimates, the GMM estimator due to Blundell and Bond (1998) is employed.

---

[77] See Sutton (1997).
[78] See ECON Report 25 (2007).

According to the results, univariate tests of Gibrat's Law fail to reject the null hypothesis that growth is random, suggesting no evidence of a systematic variation in grow performance between large and small banks. Multivariate tests of Gibrat's law, aimed at testing the sensitivity of the "random" process to various bank-level determinants such as profitability, efficiency and differentiation, show that, although there is limited systematic variation in growth rates, e.g. more diversified banks experienced higher growth rates, the growth process indeed is essentially stochastic.

The rest of the paper is organized as follows. Chapter 2: section 2.1 is devoted to a brief discussion of stochastic firm growth, while Section 2.2 briefly reviews the empirical literature on regression-based tests of Gibrat's Law in the context of banking and finance. Chapter 3 thoroughly describes the methodology employed, while Chapter 4: sections 4.1-4.2 formulate the testable hypotheses derived from the univariate and multivariate models of firm growth, respectively. Section 4.3 describes the sample data. Chapter 5 presents the results and Chapter 6 concludes.

2. Background

2.1 Stochastic firm growth theory

To date, a large number of studies on the subject of firm growth theory have been undertaken.[79] One important part of this research embraces theories which postulate certain outcomes for firm size distribution and industrial concentration, i.e. theories on stochastic and evolutionary growth.[80] Stochastic firm growth theories emanate from Gibrat (1931), whose law of proportionate effect postulates that the proportionate growth rate of incumbent firms is completely randomly determined and hence independent of systematic factors such as initial size or previous growth rates. In other words, factors that influence firm growth, such as growth of demand, managerial talent, technical innovations, organisational structure and luck, are distributed across firms in a manner which cannot be predicted from information about firm's current size or its previous growth performance.[81] It is by now well known that the implication of the LPE is a firm-size distribution which over time becomes increasingly skewed, and in the limit will approximate the log-normal

[79] Hart (2000) provides an extensive review of the theoretical and empirical literature on firm growth.
[80] This literature is thoroughly reviewed in Sutton (1997).
[81] Cf. Goodard *et al*. (2001).

distribution.[82] Thus the industry will tend to become more concentrated and dominated by a handful of large firms, even in the absence of systematic factors that would enable large firms to grow faster, such as scale economies or superior X-efficiency.

While Gibrat's theory had little immediate impact, the 1950s and 1960s saw a revival of stochastic firm growth theory.[83] The new models retained the law to specify the size-growth relationship for surviving firms, but elaborated in particular on the assumptions made about entry and exit and their role in influencing the size-growth relationship.[84]

Meanwhile, a growing number of empirical tests of the LPE were conducted, which can be grouped into two main categorical approaches: The first category comprises studies which base their analysis on empirical firm size distributions, and which tests for the LPE using goodness-of-fit tests.[85] The second category consists of studies which examine the size-growth relationship using more direct tests, based on regression analysis.[86] The advantage of regression-based tests is that they are able to incorporate dynamic influences on firm size distributions, such as persistence of growth or heteroscedasticity in growth rates. Hart (1962) identifies the following implications of the LPE: (1) large and small firms should have the same average proportional growth; (2) no heteroscedasticity in growth rates; (3) the firm size distribution should be log-normal; and (4) the relative dispersion of firm sizes should increase over time. Using these properties as a basis for regression-based tests, he

---

[82] The formal demonstration below is based on Steindl (1965) and reproduced in Sutton (1997, 1998): Let x_t and ε_t denote the size of a firm at time t and a random variable capturing the proportionate rate of growth between (t-1) and t so that:

$$x_t - x_{t-1} = \varepsilon_t x_{t-1} \qquad \text{[i]}$$

implying

$$x_t = (1+\varepsilon_t)x_{t-1} = x_0(1+\varepsilon_1)(1+\varepsilon_2)\cdots(1+\varepsilon_t). \qquad \text{[ii]}$$

Now, for short time intervals, it is reasonable to consider ε_t as small, justifying the approximation $\ln(1+\varepsilon_t) \approx \varepsilon_t$. Thus, taking logs, condition [ii] becomes:

$$\ln x_t \approx \ln x_0 + \varepsilon_1 + \varepsilon_2 + ... + \varepsilon_t. \qquad \text{[iii]}$$

By assuming the increments ε_t to be independent variates with mean *m* and variance σ^2, we have that as $t \to \infty$, the term $\ln x_0$ will be small compared to $\ln x_t$, so that $\ln x_t$ is approximated by a normal distribution with mean *mt* and variance $\sigma^2 t$. In other words, the limiting distribution is lognormal.

[83] See Steindl, (1965) for a review.
[84] See Hart and Prais (1956) and Ijiri and Simon (1977).
[85] See Simon and Bonnini (1958); Steindl (1965); and Ijiri and Simon (1977).
[86] See e.g. Hart (1962) and Mansfield (1962).

finds no evidence against the LPE for various industries during the 1930s, 1940s and 1950s.

According to Sutton (1997), the contribution of Mansfield (1962) is of particular interest. Mansfield points out that the previous inconclusive findings about the validity of the LPE emanates from using three different types of samples: (1) all firms (including those that fail to survive during the period); (2) surviving firms only; and (3) well-established firms, i.e. firms that have reached the minimum efficient scale of operation (MES), and thus have exhausted economies of scale.[87] Overall, Mansfield concludes that smaller firms have higher and more variable growth rates than larger firms, while there is support for the LPE for firms operating above the MES.

In the late 1970s and 1980s, following a revival of empirical work in the area, a number of economic, i.e. not entirely stochastic, models of firm growth were introduced. These models introduced stochastic elements into conventional maximizing models[88] In Jovanovic (1982), each firm's cost curve is subjected to randomly-distributed firm-specific shocks. Over time a firm learns about the effects of these shocks on its efficiency. Firms experiencing favourable shocks grow and survive. Others do not grow and may decline and even leave the industry (Hart, 2000). Accordingly, larger firms are likely to be older than smaller firms, since they have benefit from learning economies of scale, enabling them to avoid making costly mistakes. The implication is that large firms' growth is subject to less variation than that of smaller firms.

More recently, Sutton (1998) argues for the need of an integrated theory capable of explaining variations between industries in concentration and in the shape of their firm size distributions. Markets that tend to fragment into separate submarkets remain less concentrated than those that tend to remain homogeneous. For any given concentration ratio, firm size distributions are modelled as the outcome of a dynamic process in which there is a fixed probability that any submarket will be contested by an entrant, which may be either an established firm operating in other submarkets or a new firm. It is possible to derive a theoretical firm size distribution that would apply if the probability of successfully contesting a new submarket were the same for established and new firms. In Sutton's framework, this case is analogous to the LPE. Departures from this theoretical size distribution would occur if established firms

[87] MES is defined as the output level at which a firms average cost curve stops falling.
[88] See e.g. Jovanovic (1982) and Sutton (1991,1997,1998).

enjoyed advantages over new firms, affording a higher probability of successfully contesting new submarkets as they arise (cf. Goddard *et al.* 2001).

Nelson and Winter (1982) propose an evolutionary model of firm growth. The evolutionary approach to firm growth implies that there is some serial correlation in growth: "success breeds success and failure breeds failure". Thus this is in contrast to purely stochastic models of growth, such as the LPE, which postulate that the proportionate growth of surviving firms is random and hence independent of previous success (cf. Hart, 2000). Nelson and Winter (1982) avoids strict maximizing assumptions in favour of weaker rationality assumptions, and raises some fundamental questions as to the appropriateness of making strong rationality and informational assumptions on agents who face continuing technological change (cf. Sutton (1998), p. 244). Instead of optimising, agents tend to react automatically to changes in the market environment using routines which are specific to the firm. Successful routines which have produced growth in the past are likely to do so in the future. It is true that circumstances change, but successful firms have successful routines for changing previous methods to meet new market environments (cf. Hart, 2000).

2.2 The Law of proportionate effect and banking - A review of the literature

The majority of previous empirical tests of the LPE have been based on cross-sectional regressions of logarithmic growth over a certain time interval on initial log size, sometimes (more recently) including a term accounting for persistency of growth. The first researchers who tested for the LPE using banking data were Alhadeff & Alhadeff (1964), who compare growth rates for the 200 largest banks to the average of the whole sample, and obtained that the group of large banks tended to grow more slowly than average during 1930-1960. Rhoades & Yeats (1974) deal with a sample of 600 US banks for the period 1960-71, where they also were able to distinguish between internal and external (merger) growth. The conclusion reached was that the group of medium-sized banks experienced the highest internal growth.

Tschoegl (1983) proposes three tests relating to the LPE: (1) growth is independent of initial size; (2) growth does not persist from one period to the next, and (3) the variability of growth between banks is independent of initial size. The acceptance of these three hypotheses implies that concentration will increase over time, and that the LPE is valid in its strongest form. Based on a sample of the largest international banks which were in operation in 1969-1977, Tschoegl (1983) obtains no significant size-growth relationship. However, the variability of growth declines with size, indicating that smaller banks exhibit more variable growth rates than larger banks, contradicting the LPE in its strongest form. Finally, there is evidence of insignificantly positive persistence of growth.

Vennet (2001) investigates growth patterns for the aggregate bank sectors in 23 OECD countries, including Sweden, for the time span 1985-94. For the sub-period 1985-89, size is mean-reverting, implying that smaller bank sectors were catching up their larger counterparts. Factors such as enlarged access to revenue sources, growing internationalization of trade in financial services and increased competition were suggested as explanations for this finding. By contrast, the results for the period 1990-94 support the LPE, indicating that the largest banks were reclaiming their dominance over world banking. Moreover, the results are robust to whether total assets or adjusted total assets (total assets plus off-balance-sheet activities) are used to measure size.

Wilson and Williams (2000), test for the LPE using a sample of European banks. The study covers the period from 1990 to 1996. For the purposes of robustness, several definitions of size are considered, i.e. total assets, equity and off-balance sheet activities). Apart from Italy, no significant size-growth relationship is obtained. Large banks are found to have less variable growth rates (in line with Tschoegl, 1983), suggesting that large banks enjoy diversification advantages (Singh and Whittington, 1968) or that they are able to benefit from learning economies of scale (Jovanovic, 1982).

Hardwick and Adams (2002) examine the relationship between size and growth in the UK life insurance industry, using 1987-1996 data in a multivariate setting. With regard to the whole period, no significant size-growth relationship is obtained. However, more diversified life insurance firms experienced higher growth rates on average than more specialized life insurers. Other firm-specific determinants of growth, such as profitability or cost efficiency, turns out to be insignificant.

Goddard *et al*. (2002) investigate the size-growth relationship of US credit unions during the 1990s, using univariate and multivariate cross-sectional and panel estimation techniques. In general, larger credit unions are found to grow faster than their smaller counterparts. Moreover, they identify several systematic factors explaining why larger firms grow faster, and conclude that growth is not randomly driven but highly systematic.

Goddard *et al.* (2004) investigate firm-growth dynamics over the period 1992-1998 for a sample of commercial, savings, and cooperative banks drawn from five major European countries. The results reveal little or no evidence of mean-reversion in bank sizes. Banks maintaining a high capital-asset ratio are found to grow more slowly, and growth is linked to macroeconomic conditions. Otherwise, there appear to be few if any other factors that have a strong or systematic influence on bank growth.

3. The methodology

The relationship between firm growth and initial firm size is estimated using a multivariate dynamic panel data (DPD) model with individual effects. The model is of the form:

$$\Delta y_{i,t} = (\alpha - 1)y_{i,t-1} + \mathbf{x}'_{i,t}\boldsymbol{\beta} + (\eta_i + v_{i,t}) \qquad [3.1]$$

for $i= 1,\ldots,N$; $t = 2,3,\ldots,T$. $y_{i,t}$ is the logarithmic size of bank i in period t; $\Delta y_{i,t} \equiv y_{i,t} - y_{i,t-1}$ is logarithmic growth of bank i between t-1 and t; α is the autoregressive coefficient; and $\boldsymbol{\beta}$ a vector of parameters, corresponding to the vector of contemporaneous and one-lagged explanatory variables $\mathbf{x}_{i,t}$. The variables in $\mathbf{x}_{i,t}$ may be exogenous, predetermined or endogenous with respect to the disturbance term $v_{i,t}$; assumed to be independent across individuals and serially uncorrelated. It is also assumed that $E(v_{i,t}) = 0$ and $\text{var}(v_{i,t}) = \sigma_v^2 > 0$. Finally, η_i denotes a stochastic individual bank level effect that captures time-invariant unobserved heterogeneity. It is assumed that $E(\eta_i) = \mu_\eta$ and $\text{var}(\eta_i) = \sigma_\eta^2$.

In DPD models with unobserved individual effects, the Ordinary Least Squares (OLS) estimator is known to be inconsistent, due to the correlation between the lagged dependent variable $y_{i,t-1}$ and the individual effects η_i. Moreover, this correlation and hence the inconsistency of the estimator still persists in panels where N or $T \to \infty$. Standard results for omitted variable bias indicate that, at least in large samples, the OLS estimator is biased upwards (cf. Bond, 2002).

The usual panel estimators are also inappropriate. For instance, the within groups (fixed effects) estimator removes this source of inconsistency by transforming the equation to eliminate η_i. However if T is small and fixed, this transformation induces a non-negligible correlation between the transformed lagged dependent variable and the transformed error term. The order of the bias is $1/T$ and its consistency depends on T being large. Monte Carlo simulations have shown that this bias can be a serious problem even when T=30, in particular when α is high (see e.g. Blundell and Bond, 1998). Standard results for omitted variables bias indicate a downward bias, at least in large samples (Bond, 2002). The random effects GLS estimator is also biased in a DPD model with individual effects.

However, various consistent estimators are available (see e.g. Baltagi, 2001). Here I employ the frequently used instrumental variable GMM[89] (generalized method of moments) estimator proposed by Arellano and Bond (1991). To get rid of the fixed bank-level effects, equation [3.1] is transformed using e.g. first differencing (Arellano and Bond, 1991) or forward orthogonal deviations[90] (Arellano and Bover, 1995):

$$\Delta\bar{y}_{i,t} = (\alpha - 1)\bar{y}_{i,t-1} + \bar{\mathbf{x}}'_{i,t}\boldsymbol{\beta} + \bar{v}_{i,t} \quad i = 1, \ldots, N \qquad \text{for } t = 3, \ldots, T \qquad [3.2]$$

where the bar indicates transformed variables.

The resulting correlation between the lagged dependent variable $\Delta y_{i,t-1}$ and the transformed error term $\bar{v}_{i,t}$ necessitates the use of instrumental variables estimation. The Arellano and Bond (1991) differenced GMM(DIF) estimator utilizes the set of orthogonal moment conditions that exist between lagged *levels* of the dependent and independent variables used as instruments and the set of the transformed equations [3.2].

The moment conditions available depends on what is assumed about the correlation between $\mathbf{x}_{i,t}$ and η_i, as well as between $\mathbf{x}_{i,t}$ and $v_{i,t}$. I assume here that all the explanatory variables in $\mathbf{x}_{i,t}$ (besides $y_{i,t}$) are potentially correlated with η_i. With

[89] See Hansen (1982).

[90] The orthogonal deviations transformation (Arellano and Bover, 1995) is an alternative to first-differencing, which involves first-differencing followed by a GLS transformation to remove the resulting serial correlation induced by first-differencing. Thus if the original errors are uncorrelated, so are the transformed errors. Although first-differencing and orthogonal deviations generate quite similar parameter estimates, the latter method has been shown to offer superior efficiency in models with predetermined variables (Maeshiro and Vali, 1988). Formally, the transformation involves subtracting from each observation the average of future observations in the sample for the same individual, followed by a weighting to standardize the variances:

$$\bar{x}_{i,t} = \left(x_{i,t} - \frac{x_{i,t+1} + \ldots + x_{i,T}}{T - t} \right)\left(\frac{T - t}{T - t + 1} \right)^{1/2} \quad \text{for } t = 1, \ldots, T - 1.$$

where $\bar{x}_{i,t}$ denotes a transformed dependent or independent variable.

respect to $v_{i,t}$, a variable $x_{i,t}$ may be (1) endogenous in the sense that $E(x_{i,t}v_{i,s}) \neq 0$ for $s \leq t$ and equal to zero otherwise; (2) predetermined in the sense that $E(x_{i,t}v_{i,s}) \neq 0$ for $s < t$ and zero otherwise; or (3) strictly exogenous in the sense that $E(x_{i,t}v_{i,s}) = 0$ for all t and s.[91]

Under the weakest assumption (1), the variables in $\mathbf{x}_{i,t}$ are treated in the same way as the dependent variable $y_{i,t}$. This means that besides the vector $\{y_{i1}, y_{i2}, ..., y_{i,T-2}\}$, second and higher-order lags of each variable in $\mathbf{x}_{i,t}$ are available as valid instruments in the transformed equation for t = 3, 4,…,T. The corresponding instrument matrix is a block diagonal matrix of the form:

$$\mathbf{Z}_i = \begin{bmatrix} \{y_{i1}, \mathbf{x}'_{i1}\} & & & 0 \\ & \{y_{i1}, y_{i2}, \mathbf{x}'_{i1}, \mathbf{x}'_{i2}\} & & \\ & & \ddots & \\ 0 & & & \{y_{i1}, ..., y_{is}, \mathbf{x}'_{i1}, ..., \mathbf{x}'_{is}\} \end{bmatrix} \quad \text{for } s = 1,\ldots, T - 2. \quad [3.3]$$

Under the stronger assumption (2), there is no contemporaneous correlation between $x_{i,t}$ and $v_{i,t}$, implying that $x_{i,t-1}$ is available as an additional valid instrument. The enlarged instrument matrix is given by:

$$\mathbf{Z}_i = diag\{y_{i1},..., y_{i,s}, \mathbf{x}_{i1},...\mathbf{x}_{i,s+1}\} \qquad \text{for } s = 1, \ldots T - 2. \qquad [3.4]$$

Finally, under the strict exogeneity assumption (3), all the x_i's in the vector $\mathbf{x}'_i = (x_{i1}, x_{i2},..., x_{iT})$ are valid instruments in each of the transformed equations, corresponding to a further enlarged instrument matrix given by:

$$\mathbf{Z}_i = diag\{y_{i1},..., y_{i,s}, \mathbf{x}_{i1},...\mathbf{x}_{i,T}\} \qquad \text{for } s = 1, \ldots T - 2. \qquad [3.5]$$

In each case, the set of moment conditions can be compactly written as:

$$E(\mathbf{Z}'_i \overline{\mathbf{v}}_i) = 0 \qquad [3.6]$$

where $\overline{\mathbf{v}}_i = (\overline{v}_{i3}, \overline{v}_{i4},..., \overline{v}_{iT})'$.

The GMM(DIF) estimator is computed by minimizing a weighted quadratic form of the corresponding sample moments $\left(N^{-1}\sum_{i=1}^{N} \mathbf{Z}'_i \overline{\mathbf{v}}_i\right)$:

[91]Obviously, $x_{i,t}$ may consist of a combination of endogenous/predetermined/exogenous variables. In the regressions below, I make the simplifying assumption that all explanatory variables are either endogenous, predetermined or strictly exogenous, and subsequently test the validity of each assumption using a difference Sargan test of overidentifying restrictions.

$$\hat{\delta}_{GMM} = \arg\min_{\delta} \left\{ \left(N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i' \overline{\mathbf{v}}_i \right)' \mathbf{W}_N \left(N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i' \overline{\mathbf{v}}_i \right) \right\} \qquad [3.7]$$

where $\mathbf{W}_N$ is a positive definite weighting matrix. The optimal choice for $\mathbf{W}_N$ is computed as

$$\mathbf{W}_N = \left(N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i' \hat{\overline{\mathbf{v}}}_i \hat{\overline{\mathbf{v}}}_i' \mathbf{Z}_i \right)^{-1} \qquad [3.8]$$

where $\hat{\overline{\mathbf{v}}}_i$ contains residuals from a preliminary consistent estimator. Using [3.8] yields the two-step GMM estimator. To obtain a preliminary (one-step) estimator, I follow Arellano and Bond (1991) and use $\mathbf{W}_{1N} = \left(N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i' \mathbf{H} \mathbf{Z}_i \right)^{-1}$, where $\mathbf{H}$ is a $(T-2)\times(T-2)$ matrix with 2's on the main diagonal, -1's on the first off-diagonals and zeros elsewhere. The one-step and two-step estimators are asymptotically equivalent if the $v_{i,t}$ are homoscedastic and not correlated. The two-step GMM(DIF) estimator may be more efficient in the presence of heteroscedasticity, although simulation studies have indicated either very limited efficiency gains (cf. Bond, 2002) or even efficiency losses (e.g. Arellano and Bond (1991), Kiviet (1995) and Judson and Owen (1996)) compared to the one-step estimator.

The solution to [3.7] is given by:

$$\hat{\boldsymbol{\delta}}_{GMM} = \left[\left(\sum_{i=1}^{N} \overline{\mathbf{X}}_i' \mathbf{Z}_i \right) \mathbf{W}_N \left(\sum_{i=1}^{N} \mathbf{Z}_i' \overline{\mathbf{X}}_i \right) \right]^{-1} \left(\sum_{i=1}^{N} \overline{\mathbf{X}}_i' \mathbf{Z}_i \right) \mathbf{W}_N \left(\sum_{i=1}^{N} \mathbf{Z}_i' \overline{\mathbf{y}}_i \right) \qquad [3.9]$$

or in compact notation as:

$$\hat{\boldsymbol{\delta}}_{GMM} = \left(\overline{\mathbf{X}}' \mathbf{Z} \mathbf{W}_N \mathbf{Z}' \overline{\mathbf{X}} \right)^{-1} \overline{\mathbf{X}}' \mathbf{Z} \mathbf{W}_N \mathbf{Z}' \overline{\mathbf{y}} \qquad [3.10]$$

where $\hat{\boldsymbol{\delta}}_{GMM}$ contains the GMM estimates of α and $\boldsymbol{\beta}$ in [3.1]; $\overline{\mathbf{X}}$ denotes a matrix of observations on the transformed regressors (including the lagged dependent variable); and $\overline{\mathbf{y}}$ contains transformations of the dependent variable $\Delta y_{i,t}$.

The GMM(DIF) estimator is subject to a severe weak instrument bias when the individual series are highly persistent over time, and the number of time-series observations is moderately small (see Blundell and Bond (1998); Blundell *et al.*, (2000)). If the variables are highly persistent, lagged levels of these variables used as instruments are only weakly correlated with subsequent first-differences of the same variables and therefore constitute weak instruments. Simulations have shown that in particular the GMM(DIF) estimator suffers from large finite sample downward bias and very low precision when the instruments used are weak (cf. Blundell *et al*., (2000)).

The extended system GMM estimator (Blundell and Bond, 1998) combines the moment conditions exploited in the GMM(DIF) estimator with additional moment conditions for the untransformed level equations as suggested by Arellano and Bover (1995). Both simulations and empirical experience have shown that the system (GMM(SYS)) estimator has much better finite sample properties (less bias and greater precision) than the GMM(DIF) estimator in the presence of persistent time-series (see Blundell *et al*., 2000 and Bond, 2002).

The GMM(SYS) estimation procedure uses the same matrix of instruments for the transformed equations as above. Assuming (1) that the initial conditions satisfy mean stationarity, implying $E(\Delta y_{i2}\eta_i) = 0$ and (2) that $\Delta x_{i,t}$ are uncorrelated with η_i, the following additional non-redundant moment conditions that remain informative in the case of persistent series can be exploited:

$$E(\Delta z_{i,t-s}(\eta_i + v_{i,t})) = 0 \qquad \text{for } s = 1;\ t = 3, \ldots, T \qquad [3.11]$$

where $z_{i,t}$ denotes any variable in $\{y_{i,t}, \mathbf{x}'_{i,t}\}$. These additional moment conditions allow us to use suitable lagged first-differences of the variables as instruments for the level equations. Given that the moment conditions in [3.6] are exploited, only the most recent difference is used as an instrument in the levels specification (s = 1). Using further lagged differences would result in redundant moment conditions (Arellano and Bover, 1995).

The GMM(SYS) estimator is thus based on a stacked system of (1) the transformed equations, and (2) the level equations, with the moment conditions in [3.6] applied to the first part and those in [3.11] applied to the second part (cf. Blundell and bond, 1998). The extended instrument matrix used is of the form:

$$\mathbf{Z}_i^+ = \begin{bmatrix} \mathbf{Z}_i & & & & 0 \\ & \Delta y_{i1}, \Delta\mathbf{x}'_{i2} & & & \\ & & \Delta y_{i2}, \Delta\mathbf{x}'_{i3} & & \\ & & & \ddots & \\ 0 & & & & \Delta y_{i,s}\ \Delta\mathbf{x}'_{i,s+1} \end{bmatrix} \qquad \text{for } s = 1, \ldots\ T-2. \qquad [3.12]$$

Estimation then proceeds in two steps, as with the GMM(DIF) estimator (Blundell and Bond, 1998).

The consistency of the GMM estimators depends crucially on the validity of the instruments. No serial correlation in $v_{i,t}$ is the key identifying assumption with respect to instrument validity. The validity of the instruments can be examined by testing whether the transformed (differenced) error term is second-order serially correlated. The test statistics (see Arellano and Bond, 1991) are asymptotically standard normally distributed under the null of no second-order serial correlation. An additional test of instrument validity is the Sargan test which is applicable in overidentified models. The hypothesis being tested is that the instruments are uncorrelated with the errors in the transformed equations, and hence acceptable. Under the null, the test statistic (see e.g. Arellano and Bond, 1991) is asymptotically distributed as a chi-square with degrees of freedom equal to the number of instruments minus the number of parameters.

The validity of the additional moment conditions [3.11] is tested using the difference-Sargan test proposed by Arellano and Bond (1991). Since $\mathbf{Z}_i$ is a strict subset of $\mathbf{Z}_i^+$, the difference-Sargan test statistic is computed by subtracting the Sargan statistic for the GMM(DIF) estimator from that of the GMM(SYS) estimator. The test-statistic is asymptotically chi-square distributed under the null hypothesis of validity of the extra instruments, with degrees of freedom equal to the difference in degrees of freedom of the GMM(SYS) estimator and the GMM(DIF) estimator.

4. Data and testable hypotheses

This chapter is divided into three sections. Sections 4.1-4.2 formulate the testable hypotheses derived from the univariate and multivariate models of firm growth, respectively. Subsequently, Section 4.3 describes the sample data.

4.1 Univariate growth model

The univariate test of the LPE (Gibrat's random walk growth hypothesis) is based on the following model of firm growth:

$$\Delta y_{i,t} = (\alpha - 1) y_{i,t-1} + \rho \Delta y_{i,t-1} + \boldsymbol{\delta}_t + \eta_i + \varepsilon_{i,t} \qquad [4.1]$$

where $y_{i,t}$ denotes the size of bank *i* in period *t*, measured as the natural logarithm of total assets.[92] Notations are the same as in Chapter 3. In addition, ρ denotes a parameter capturing serial correlation in growth rates (persistence of growth); $\boldsymbol{\delta}_t$ denotes a vector of time effects; and $\varepsilon_{i,t}$ is an error term, normally and i.i.d. distributed under the null of no size-growth relationship, i.e. growth follows a random walk ($\boldsymbol{H_0}$: $\alpha = 1$).[93]

Based on the univariate model two of Tschoegl's (1983) testable propositions with respect to the LPE (cf. Section 2.2) are directly testable: (1) growth is independent of initial size and (2) high or low growth in one period does not persist to the next. Testing hypotheses (1) amounts to test the null hypothesis that $\alpha = 1$, implying that growth is non-explosive and unrelated to size. Failure to reject this hypothesis is consistent with the LPE. By contrast $\alpha < 1$ implies that firm sizes are mean-reverting. Small banks tend to grow faster than large banks, suggesting that over time, the size of all banks is reverting towards some long-run mean value, and there is no tendency for industrial concentration to increase. Under mean-reversion it is assumed that η_i >0 and the average size to which banks tend to revert back to is equal to $\eta_i/(1-\alpha)$. If $\eta_i = \eta$ ($\sigma_\eta^2 = 0$), there is a common long-term mean size for all banks, while if $\eta_i \neq \eta$ ($\sigma_\eta^2 > 0$) there are heterogeneous, bank-specific long-term values. Finally $\alpha > 1$ implies growth trajectories that are explosive implying rapidly increased size dispersion. Banks grow proportionately faster as they get larger. This can go on for a finite period but is unlikely to last for long. Note that when $\alpha \geq 1$, it is assumed that $\eta_i = \eta = 0$, and η_i has no interpretation in terms of mean-reversion.[94]

---

[92] Ever since Tschoegl (1983), total assets and/or total equity have frequently been used as measures of size in tests of the LPE within the context of banking and finance (see e.g. Wilson and Williams (2000); Vennet (2001); Hardwick and Adams (2002); Goddard, McKillop and Wilson (2002); and Goddard, Molyneux and Wilson (2004)). A few recent tests account for the shift in banking activities towards increased engagement in off-balance-sheet business activities (see e.g. Vennet (2001) and Goddard *et al*. (2004)).

[93] It is assumed that the true data generating process is given by:

$$\Delta y_{i,t} = (\alpha - 1)y_{i,t-1} + \boldsymbol{\delta}_t + \eta_i + \mu_{i,t}; \quad \mu_{i,t} = \rho\mu_{i,t-1} + v_{i,t} \quad \text{(i)}$$

For the purposes of panel estimation, (i) can be re-writtten as follows (cf. Goddard, McKillop and Wilson, 2002):

$$\Delta y_{i,t} = (\alpha - 1)y_{i,t-1} + \rho\Delta y_{i,t-1} + \boldsymbol{\delta}_t + \eta_i + \varepsilon_{i,t} \qquad \text{(ii)}$$

where $\varepsilon_{i,t} = v_{i,t} + \rho(1-\alpha)y_{i,t-2}$, so that $\varepsilon_{i,t} = v_{i,t}$ under $\boldsymbol{H_0}$.

[94] Assuming $\eta_i \neq 0$ would allow for a deterministic trend specific to each bank, which could exist but which would be very difficult to identify unless the number of observations per bank is quite large. The

Testing proposition (2) amounts to test the null hypothesis that ρ =0. If ρ >0 (ρ <0) there is positive (negative) persistence in growth rates. Previous studies have reported both positive and negative estimates of ρ (Goddard *et al*., 2004). If ρ >0 under the hypothesis of random walk growth ($\boldsymbol{H_0}$: $\alpha = 1$), this means that banks that are fortunate in drawing an above-average growth rate in one year have a relatively greater chance of repeating this successful performance in the following year: success breeds success and failure breeds failure (cf. Chapter 2). The implication of such a growth process is that average bank size tend to increase at a faster rate than if ρ =0. Industry concentration also tends to increase at a faster rate (see Goddard *et al*., 2001).

Failure to reject the null hypotheses of both tests implies that the univariate model cannot reject the LPE.[95]

4.2 Multivariate growth model

Recently it has been noticed that the univariate test of the LPE might be overly simplistic in the sense that it cannot distinguish between, on the one hand, true random growth and, on the other hand, growth that appears to be random due to an aggregation of various possible determinants of firm growth (cf. Geroski *et al*., 1997). Consequently, more recent empirical tests of the LPE have increasingly begun to test the sensitivity of the "random" growth process to a range of firm- and industry-level growth determinants within a multivariate structural model framework.

The issue of firm growth has been addressed extensively within manufacturing. Within manufacturing, the stylized facts of the empirical literature are that (1) firm sizes are mean-reverting, which contradicts to the LPE, while (2) for sub-samples of large and well-established firms, the LPE tends to be confirmed (Gibrat's legacy).[96]

Much less attention has been paid to the determinants of growth within service industries such as banking. The literature so far includes Vennet (2001), Goddard, McKillop and Wilson (2002), and Goddard *et al*. (2004) who points out that "departures from the LPE in banking might be explained by factors including efficiency variations between large and small firms, the adoption of entry-deterring

---

possibility of a common deterministic trend is captured, however, through the time effects, δ_t (Goddard, McKillop and Wilson, 2002).

[95] The third testable proposition formulated by Tschoegl (1983), i.e. that growth between banks is independent of initial size can be examined by performing a Lagrange multiplier test of the residuals of the estimated equations (see e.g. Goddard *et al*., 2001).

[96] See Sutton (1997).

strategies or the exercise of market power by large firms, or the superior flexibility and innovativeness of small firms".

In order to examine if the growth process is sensitive to bank-level performance indicators such as operating efficiency and profitability, I test for the LPE based on the following multivariate firm growth model:

$$\Delta y_{i,t} = (\alpha - 1) y_{i,t-1} + \rho \Delta y_{i,t-1} + \mathbf{x}'_{i,t}\boldsymbol{\beta} + \boldsymbol{\delta}_t + \eta_i + \varepsilon_{i,t} \qquad [4.2]$$

where $\mathbf{x}_{i,t}$ denotes a vector of current and one-year lagged bank-level variables that control for differences in performance among the sample banks. Definitions of the control variables are shown in Table 4.1.[97]

Table 4.1: Bank performance variables

Regressor	Definition
ROA	Return on total assets = EBIT/Total assets
EFF	Cost-income ratio = Total operating costs / total revenues
MIX	Revenue Mix = Total non-interest revenues / Total revenues
TYPE	Bank type = 1 for commercial banks; 0 for savings banks

As a measure of profitability, I consider return on total assets, measured by EBIT (earnings before interest and taxes) divided by total assets. The linkage between profitability and firm growth is ambiguous: On the one hand, retained profits constitute a principal source of capital. Accordingly, more profitable firms should grow faster as they have more finance available. In addition, a higher rate of return should *per se* act as an incentive to make new investments. On the other hand, banks may trade current profits for future growth if managers have discretion to pursue their own objectives, which perhaps are more aligned with firm size than with profit maximization.[98] If indeed growth and profit are competing objectives, we can expect an inverse relationship between current profit rate and growth. By contrast, the coefficient of lagged *ROA* is expected to be positive as retained profits are important for future growth.

[97] All these variables have been considered in previous studies. See e.g. Vennet (2001); Hardwick and Adams (2002); Goddard, McKillop and Wilson (2002).
[98] See e.g. Goddard et al. (2004) and the cited references therein for more on managerial theories and empirical evidence.

The cost-income ratio is used as a rough proxy for operating efficiency, measured as total operating costs (total costs net of interest expenses) divided by total revenues. A high-cost ratio indicates suboptimal performance which may be due to technical inefficiency (X-inefficiency), allocative inefficiency or un-exploited economies of scale and scope.[99] A high cost ratio is expected to impact negatively on growth prospects, all else equal.[100]

MIX controls for variation in diversification. In view of the fact that banks are multi product firms, a possible source of inefficiency is the choice of a suboptimal mix of products. In the presence of scope economies, a more diversified bank is expected to operate more efficiently and thus be able to grow faster.

Finally, it seems reasonable to expect that ownership differences have implications for the growth prospects of banks. Savings banks pursue different goals than pure profit maximization. Thus they can be expected to accumulate less capital than commercial banks over time, and therefore grow more slowly. In addition, their growth opportunities might be restricted by their more limited business mix (Goddard *et al.*, 2004).

Based on these predictions, I formulate some additional propositions, which are summarized in Table 4.2.

4.3 The sample

The sample includes banks for which complete annual data were available for all years from 1995 to 2002.[101] Included among these banks are (1) nationwide commercial banks, (2) young internet banks, (3) converted savings banks, and (4) savings banks. Hence the sample represents a heterogeneous cross-section of banks, which is reflected, *inter alia*, in the large variation in size among the sample banks, as shown in Table 4.3.[102] The figures appearing in Table 4.3 are calculated from selected balance sheet and profit/loss account ex post items, and expressed in 1995 values (using the GDP deflator).

---

[99] See Berger and Humphrey (1997).
[100] Overall, technological changes such as internet banking have led to a substantial reduction in banks' cost-to-income ratios (cf. Chapter 1)
[101] On this ground, all foreign banks (branches or subsidiaries) or banks that entered or exited during this period were excluded. In order to focus on organic growth, banks that were engaged in M&As were likewise excluded (with the exception of Swedbank AB).
[102] In fact, the size distribution is highly skewed with many small banks and a few giants, as predicted by the LPE. The giants are Nordea, SEB, Svenska Handelsbanken and Swedbank. Note that the present study ignores the non-domestic business activities of these banks.

Table 4.2 Propositions with regard to firm growth in banking

P1: There is no size-growth relationship, consistent with the LPE.
P2: There is no persistence in growth rates, consistent with the LPE.
P3: Managers trade current profit for future growth. Accordingly, current growth and current profit rate are negatively associated.
P4: The higher the lagged profit rate, the more capital is available for new investments, spurring growth. Current growth and one-lagged profit rate are thus positively associated.
P5: The cost-income ratio and growth are negatively associated.
P6: More diversified banks grow faster than less diversified ones. Thus, revenue mix and growth are positively associated.
P7: Savings banks grow at a slower rate than commercial banks.

All estimations are based on panel data for N=79 and T=5[103], hence 395 observations. All reported panel data estimates are computed using DPD for OX (see Doornik et al., 2002).

Table 4.3: Descriptive Sample statistics

Variable	Mean	Std. Dev	Min	Max
SIZE	26 739	115 648	24.6	945 544
GRW	0.05508	0.1180	-0.4342	0.8977
ROA	0.01958	0.01194	-0.04892	0.08260
EFF	0.4428	0.1441	0.1224	1.096
MIX	0.1259	0.06589	0.00000	0.4763
TYPE	0.0896	0.2858	0.00000	1.000

Notes: All figures are calculated using data supplied by Statistics Sweden and the Swedish Riksbank. Size is total assets in millions of SEK.

5. Estimation and results

Before estimating the firm growth model, I follow the advice in Bond (2002) and examine the time series properties of the individual series, to mitigate any potential

[103] Three time-series observations are lost due to the inclusion of lagged variables and the transformation of the data (first-differencing or orthogonal deviations).

problems of weak instrument bias when estimating the firm growth model.[104] I also compare the consistent GMM estimator to alternative estimators which are known to be biased in opposite directions in DPD-models with fixed effects. The diagnostic analysis of the individual series are based on the AR(1) model:

$$y_{i,t} = \alpha y_{i,t-1} + (\eta_i + v_{i,t}) \qquad [5.1]$$

where $y_{i,t}$ denotes any variable and $v_{i,t}$ is serially uncorrelated. Table 5.1 presents the results. Overall, the results obtained mimic those of simulation studies aimed at investigating the finite sample properties of different estimators in AR(1) DPD models when $\alpha \rightarrow 1$.[105]

As Table 5.1 below shows, both *SIZE* and *MIX* are highly persistent. Consequently, lagged levels of these variables provide weak instruments for the transformed equations (cf. Chapter 3). This is reflected in the poor performance of the GMM(DIF) estimator, which appears to be subject to severe finite sample downward bias and low precision. By contrast, the GMM(SYS) procedure generates AR(1) coefficient estimates of the persistent variables that are much more similar in magnitude to that of the OLS, and much more precise. As expected, therefore, the difference-Sargan test accepts the validity of the additional moment conditions in [3.11].

In contrast to *SIZE* and *MIX*, the series of the other two variables, *ROA* and *EFF*, do not appear to have near-unit root properties, as shown. In these cases, we see that the two GMM procedures generate AR(1)-estimates that are more similar in magnitude and precision, although still different. The better performance of the GMM(DIF) estimator in these cases reflects that lagged levels become more informative as instruments for the transformed equations when α is smaller. This is confirmed by the Sargan test which now accepts the validity of the level instruments. Although *ROA* and *EFF* are relatively less persistent, the difference-Sargan test suggests that efficiency will improve if the additional moment conditions are exploited.

[104] The weak instrument bias problem due to highly persistent series is a relevant issue also in multivariate models (Blundell, Bond and Windmeijer, 2000).
[105] See Bond (2002) and the references cited therein.

Table 5.1: AR(1) Specifications for the growth determinant series (Unit root test)

	OLS Level	Within	GMM(DIF)	GMM(SYS)
Size$_{i,t-1}$	1.002*** (0.0041)	0.315** (0.0952)	0.319 (0.2171)	0.983*** (0.0182)
AR(2)	[0.305]	[0.550]	[0.723]	[0.337]
Sargan	_	_	[0.004]	[0.008]
Dif. Sargan	_	_	_	[0.706]
Roa $_{i,t-1}$	0.586*** (0.0599)	0.149** (0.0495)	0.295*** (0.0552)	0.284*** (0.0729)
AR(2)	[0.915]	[0.000]	[0.373]	[0.372]
Sargan	_	_	[0.157]	[0.300]
Dif. Sargan	_	_	_	[0.806]
Eff $_{i,t-1}$	0.883*** (0.0358)	0.340*** (0.0523)	0.376** (0.1548)	0.771*** (0.0441)
AR(2)	[0.336]	[0.010]	[0.412]	[0.619]
Sargan	_	_	[0.289]	[0.479]
Dif. Sargan	_	_	_	[0.838]
Mix $_{i,t-1}$	0.978*** (0.0366)	0.604*** (0.1345)	0.140 (0.1304)	0.882*** (0.1708)
AR(2)	[0.028]	[0.010]	[0.247]	[0.571]
Sargan	_	_	[0.004]	[0.016]
Dif. Sargan	_	_	_	[0.786]

Notes: ***,**,* denotes coefficient significantly different from zero at the 1, 5 and 10% level, respectively. Two-step coefficient estimates are reported together with t-ratios based on finite sample corrected standard errors (in brackets); p-values in square brackets.

The outcome of the diagnostic analysis of the individual series signals that the GMM(SYS) estimator is likely to be preferred when estimating the multivariate growth model [4.2].[106] However, exploiting all moment conditions in [3.6] and [3.11] may result in over-fitting bias. That is, while increasing the number of instruments improves efficiency, it also increases the bias in finite samples

[106] See Blundell *et al*. (2000).

Table 5.2: Two-step GMM- SYS estimation results

Dependent variable= $Grw_{i,t}$

	Univariate	Multivariate		
	(1)	(2)	(3)	(4)
Regressors		Exogenous	Predetermined	Endogenous
$Size_{t-1}$	-0.0192 (0.0442)	-0.0024 (0.0193)	-0.0071 0.0193	-0.0010 (0.0224)
Grw_{t-1}	-0.0054 (0.0184)	-0.0217 (0.0218)	-0.0172 (0.0227)	-0.0086 (0.0251)
Roa_t	-	-1.319 (1.300)	-2.504 (1.784)	-3.789 (4.904)
Roa_{t-1}	-	0.0405 (2.133)	0.0218 (2.239)	0.6035 (3.248)
Mix_t	-	0.5764* (0.3585)	0.7335* (0.4288)	0.3926 (0.4203)
Mix_{t-1}	-	-0.3878 (0.3673)	-0.3852 (0.3647)	-0.2234 (0.3501)
Eff_t	-	0.4247** (0.2162)	0.3374* 0.1824	0.5222 (0.3880)
Eff_{t-1}	-	-0.2934 (0.2360)	-0.3123 (0.2231)	-0.4887* (0.2773)
Constant	0.1024 (0.2722)	-0.0732 (0.1605)	0.0112 (0.1670)	-0.0005 (0.2453)
Type dummy	-	0.0461 (0.0806)	0.0571 (0.0852)	0.0360 (0.0971)
Time dummies	[0.000]	[0.000]	[0.000]	[0.000]
Wald (χ^2)	0.3893 [0.823]	15.58 [0.049]	14.44 [0.071]	12.09 [0.147]
Sargan test	[0.001] (14)	[0.885] (86)	[0.423] (71)	[0.142] (56)
AR(2) test	[0.548]	[0.508]	[0.592]	[0.928]
Dif. Sargan (Sys vs.Dif)	[0.317]	[1.000]	[0.968]	[0.650]
Dif. Sargan (regressors)	-	[1.000]	[0.989]	-
Instruments	Size(2,3);ΔSize(1,1)	Size(2,3);ΔSize(1,1) Roa(0,3);ΔRoa(0,0) Mix(0,3);ΔMix(0,0) Eff(0,3); ΔEff(0,0)	Size(2,3);ΔSize(1,1) Roa(1,3);ΔRoa(0,0) Mix(1,3);ΔMix(0,0) Eff(1,3); ΔEff(0,0)	Size(2,3);ΔSize(1,1) Roa(2,3);ΔRoa(1,1) Mix(2,3);ΔMix(1,1) Eff(2,3); ΔEff(1,1)

Notes: ***,**,* denotes coefficient significantly different from zero at the 1, 5 and 10% level, respectively. Variables are transformed using first differencing, as proposed by Arellano and Bond (1991). Results reported are estimated using robust standard errors (in parentheses). Figures in square brackets denote p-values. Wald is a test of joint significance of the independent variables (except constant and dummies). AR(2) is a N(0,1) test of no serial correlation in the residuals of the differenced model. Sargan is a test of overidentifying restrictions, asymptotically distributed as a chi-square under the null of instrument validity. The number of overidentifying restrictions is given in parentheses. The Difference Sargan test is used to test the validity of the additional momemt conditions in [3.11], as well as the assumption made about the regressors. As instruments, lagged levels of the variables are used (consistent with moment conditions [3.6]), together with lagged differences of the variables, consistent with [3.11]. For example, Size(2,3) denotes that the second and third lag of Size are included among the lagged level instruments, while ΔSize(1,1) denotes that the first lagged difference of Size is included among the lagged first-differenced instruments. No. of firms= 79; No. of observations = 395 (balanced panel).

(Judson and Owen, 1996). Because of this bias-efficiency trade-off, not all available lags are used as instruments.[107]

The main focus is on the two-step GMM(SYS) estimator.[108] Table 5.2 reports the results.

Model (1) reports estimates of the univariate growth model [4.1]. The estimated coefficient on $Size_{t-1}$ is negative, suggesting that smaller banks grew faster than larger ones. Likewise, the coefficient on lagged growth Grw_{t-1} is also negative, indicating negative persistence in growth and suggesting that periods of above-average and below-average growth tend to follow one another. However, both coefficients are far from being significant. The null hypothesis that both coefficients are equal to zero is not rejected by the Wald joint test (p-value=0.823). These findings are consistent with propositions P1 and P2 (cf. Table 4.2) and hence the LPE, while inconsistent with the view that the implementation of cost-saving technological innovations such as internet banking have systematically favoured large banks over small banks in terms of growth performance.

Models (2)-(4) report estimates of the multivariate growth equation [4.2]. Model (2) reports estimates assuming that all regressors (apart from $SIZE_{t-1}$) are strictly exogenous, while models (3)-(4) report estimates assuming that all regressors are predetermined and endogenous, respectively. Recall from Chapter 3 that (4) corresponds to the weakest assumption in terms of moment conditions specified, while (3) is stronger and (2) even stronger. Hence a difference Sargan test can be used to test the validity of the additional moment conditions under the stronger assumption (cf. Bond, 2002). As shown in the next to bottom row, the strongest assumption of exogenous covariates was clearly not rejected. Hence I focus on the estimates reported for model (2).

---

[107] Okui (2005) proposes a method for choosing the optimal number of instruments in AR(1) dynamic panel data models. When T=5 and N=100, for instance, the optimal number of lags included varies between 1 and 3, depending on the magnitude of the autoregressive parameter and the ratio var(η_i)/var(v_{it}).

[108] The gain in precision from using the two-step GMM estimator rather than the one-step GMM estimator is likely to be greater in this case than in the case of the GMM(DIF) estimator, since there is no one-step GMM estimator that is asymptotically equivalent to the two-step estimator, even in the case of i.i.d. disturbances (Blundell and Bond, 1998). However, as in the case of the GMM(DIF) estimator, the estimated asymptotic standard errors obtained from the two-step estimator can be severely downward biased in small samples (Windmeijer, 2005). Thus while the two-step coefficients reported in Table 5.2 are more efficient than the one-step counterparts, asymptotic inference based on the one-step errors, reported in Table A1, might be more reliable.

Before interpreting coefficient estimates, we should check the validity of the assumptions underlying the model. As shown in column (2), both the Sargan test of overidentifying restrictions and the AR(2) test suggest that the chosen instruments, reported in the bottom row, are valid (p-values are 0.885 and 0.504, respectively). Furthermore, the difference Sargan test suggests that the additional moment conditions exploited by the GMM(SYS) estimator are clearly valid (p-value is 1.000).

The Wald test of joint significance shows that the null hypothesis that all independent variables (apart from constant and dummies) are equal to zero is rejected at the 5% level.

As in the univariate model, the coefficient on $Size_{t-1}$ is insignificantly negative. However, the magnitude of the coefficient is even lower than in the univariate model, suggesting that the univariate model overstates the size-growth relationship due to the omission of significant explanatory variables included in the multivariate model. Likewise, the coefficient on Grw_{t-1} is insignificantly negative, although the magnitude of the coefficient is somewhat higher and precision is slightly better than in the univariate model.

As predicted by proposition P3 (cf. Table 4.2), the coefficient on Roa_t takes on a negative value with a magnitude of -1.319. The coefficient is more precisely estimated than that of $Size_{t-1}$ and Grw_{t-1} though insignificant at the 10% level. The evidence of a short-run growth-profit trade-off must be considered as, at most, limited.

Proposition P4 predicts that the coefficient on Roa_{t-1} should be positive, which indeed turn out to be the case. However, the precision of the estimate is very poor, suggesting no reliable linkage between one-lagged profit rate and current growth.

Proposition P5 states that less efficient banks (that is, banks operating with a higher cost-to-income ratio) should grow more slowly than more efficient banks. *A priori* a negative sign of the coefficient on Eff_t is expected. Therefore, the reported significantly positive coefficient is a puzzle. It is however possible that this variable captures the effect of an omitted correlated variable.

As predicted by proposition P6, the coefficient on Mix_t is significantly positive, suggesting that banks with a more diversified business mix were able to exploit economies of scope and therefore grow faster. With regard to proposition P7, the coefficient on the ownership type dummy is positive as expected. However, since the precision is very low, no reliable inference can be drawn with regard to differences in growth performance between commercial and savings banks.

The robustness of the results is checked with regard to a number of different estimators. Table A1 in the Appendix reports the results for the one-step GMM(SYS) estimator. As mentioned above, the asymptotic inference from the one-step standard errors might be more reliable. However, parameter estimates and standard errors are very similar to those of the two-step estimator.

Table A2 checks the sensitivity of the results with respect to transformation applied. That is, the model is estimated using the orthogonal deviations transformation (cf. Chapter 3) instead of first-differencing transformation. As shown, the results are largely similar in both cases.

Table A3 reports the results for the two-step GMM(DIF) estimator. As expected, the coefficient on $Size_{t-1}$ is subject to a severe downward bias, reflecting the poor finite sample properties of this estimator. Finally, for the purposes of comparison, Table A4 reports the results for the Pooled OLS estimator and the fixed effects estimator, which are inconsistent and upward (downward) biased in DPD models with fixed effects (cf. Chapter 3).

6. Conclusions

The purpose of this paper was to contribute to the understanding of firm growth dynamics in the new banking environment, by testing the validity of Gibrat's Law of Proportionate Effect on Swedish data. The point of departure in the paper was the expectation that large banks should be able to more fully exploit scale and scope economies associated with technological innovations such as internet banking, than smaller banks, and therefore grow faster. Using a panel of 79 Swedish banks over the period 1995-2002, I found no empirical evidence that large banks grew faster, nor any significant evidence that firm sizes were mean-reverting. Hence the Law could not be rejected. However, growth was not entirely random, as banks with a more diversified revenue mix experienced significantly higher growth rates than less diversified banks.

Bibliography

Alhadeff, C. and Alhadeff, D. (1964), "Growth of Large Banks 1930-1960", *Review of Economics and Statistics*, vol. 46, pp. 356-363.

Arellano, M. and Bond, S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, vol. 58, pp. 277-297.

Arellano, M. and Bover, O. (1995), "Another Look at the Instrumental Variable Estimation of Error-Components Models", Journal of Econometrics, vol. 68, pp. 29-51.

Baltagi, B.H. (2001), *Econometric Analysis of Panel Data*, Wiley.

Berger, A.N. and Humphrey, D.B. (1997), "Efficiency of Financial Institutions: International Survey and Directions for Future Research", *European Journal of Operational Research*, vol. 98, pp. 175-212.

Blundell, R. and Bond, S. (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*, vol. 87, pp. 115-143.

Blundell, R., Bond, S., and Windmeijer, F. (2000), "Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimators", *Advances in Econometrics*, vol. 15, pp. 53-91.

Bond, S. (2002), "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice", Institute for Fiscal Studies, Working Paper 02/09.

Doornik, J., Arellano, M., and Bond, S. (2002), "Panel Data Estimation Using DPD for OX"; http://www.doornik.com

ECON Report (2007), "Konkurrensen på Bankmarknaden".

Geroski, P., Machin, S., and Walters, C. (1997), "Corporate Growth and Profitability", *Journal of Industrial Economics*, vol. 45, pp. 171-189.

Gibrat, R. (1931), *Les Inégalités Economiques*, Paris: Sirey.

Goddard, J.A., Molyneux, P., and Wilson, J. (2001), *European Banking: Efficiency, Technology and Growth*, John Wiley & Sons.

Goddard, J.A., McKillop, D., and Wilson, J. (2002), "Credit Union Size and Growth: Tests of the Law of Proportionate Effect", *Journal of Banking and Finance*, vol. 26, pp. 2327-2356.

Goddard, J.A., Molyneux, P., and Wilson, J. (2004), "Dynamics of Growth and Profitability in Banking", *Journal of Money, Credit, and Banking*, vol. 36, pp. 1070-1090.

Hansen, L.P., (1982), "Large sample Properties of Generalized Method of Moments Estimators", *Econometrica*, vol. 50, pp. 1029-1054.

Hardwick, P. and Adams, M. (2002), "Firm Size and Growth in the United Kingdom Life Insurance Industry", *The Journal of Risk and Insurance*, vol. 69, pp. 577-593.

Hart, P.E. and Prais, S.J. (1956), "The Analysis of Business Concentration: A Statistical Approach", *Journal of the Royal Statistical Society*, series A, vol. 119, pp. 150-191.

Hart, P.E. (1962), "The Size and Growth of Firms", *Economica*, vol. 29, pp. 29-39.

Hart, P.E. (2000), "Theories of Firms' Growth and the Generation of Jobs", *Review of Industrial Organization*, vol. 17, pp. 229-248.

Ijiri, Y. and Simon, H. (1977), *Skew Distributions and the Sizes of Business Firms*, North-Holland, Amsterdam.

Jovanovic, B. (1982), "Selection and the Evolution of Industry", *Econometrica*, vol. 50, pp. 649-670.

Judson, R.A. and Owen, A.L. (1996), "Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists", Federal Reserve Board of Governors, Washington D.C.

Kiviet, J.F. (1995), "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models", *Journal of Econometrics*, vol. 68, pp. 53-78.

Mansfield, E. (1962), "Entry, Gibrat's Law, Innovation, and the Growth of Firms", *American Economic Review*, vol. 52, pp. 1023-1051.

Maeshiro, A. and Vali, S. (1988), "Pitfalls in the Estimation of a Differenced Model", *Journal of Business and Economic Statistics*, vol. 6, pp. 511-515.

Nelson, R. and Winter, S. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, Harward University Press.

Okui, R. (2005), "The Optimal Choice of Moments in Dynamic Panel Data Models", mimeo.

Rhoades, S. and Yeats, A. (1974), "Growth, Consolidation and Mergers in Banking", *Journal of Finance*, vol. 29, pp. 1397-1405.

Sargan, J.D. (1958), "The Estimation of Economic Relationships using Instrumental Variables", *Econometrica*, vol. 26, pp. 329-338.

Simon, H. and Bonini, C. (1958), "The Size Distribution of Business Firms", *American Economic Review*, vol. 48, pp. 607-617.

Singh, A. and Whittington, G. (1968), *Growth, Profitability and Valuation*, Cambridge University Press.

Steindl, J. (1965), *Random Processes and the Growth of Firms: A Study of the Pareto Law*, London: Griffith.

Sutton, J. (1991), *Sunk Costs and Market Structure*, London: MIT Press.

Sutton, J. (1997), "Gibrat's Legacy", *Journal of Economic Literature*, vol. 35, pp. 40-59.

Sutton, J. (1998), *Technology and Market Structure. Theory and History*, The MIT Press.

Tschoegl, A. (1983), "Size, Growth and Transnationality Among the World's Largest Banks", *Journal of Business*, vol.56, pp.187-201.

Vander Vennet, R. (2001), "The Law of Proportionate Effect and OECD Bank Sectors", *Applied Economics*, vol. 33, pp. 539-546.

Wilson, J. and Williams, J. (2000), "The Size and Growth of Banks: Evidence from Four European Countries", *Applied Economics*, vol. 32, pp. 1101-1109.

Windmeijer, F. (2005), "A Finite Sample Correction for the Variance of Linear Two-Step GMM Estimators", *Journal of Econometrics*, vol. 126, pp. 25-51.

Appendix

Table A1: One-step GMM- SYS estimation results
Dependent variable= $\boldsymbol{Grw}_{i,t}$

	Univariate	**Multivariate**		
	(5)	**(6)**	**(7)**	**(8)**
Regressors		Exogenous	Predetermined	Endogenous
$Size_{t-1}$	-0.0206 (0.0433)	-0.0034 (0.0195)	-0.0068 (0.0189)	-0.0035 (0.0198)
Grw_{t-1}	-0.0082 (0.0173)	-0.0193 (0.0213)	-0.0180 (0.0216)	-0.0088 (0.0240)
Roa_{t}	-	-1.547 (1.238)	-2.552 (1.683)	-4.010 (4.457)
Roa_{t-1}	-	-0.1646 (2.077)	-0.0836 (2.088)	0.5644 (3.129)
Mix_{t}	-	0.6350* (0.3460)	0.7466* (0.4253)	0.3802 (0.3546)
Mix_{t-1}	-	-0.3779 (0.3733)	-0.3784 (0.3760)	-0.1475 (0.2873)
Eff_{t}	-	0.4124* (0.2141)	0.3441* (0.1921)	0.4887 (0.3556)
Eff_{t-1}	-	-0.2956 (0.2254)	-0.3185 (0.2273)	-0.4625* (0.2545)
Constant	0.1173 (0.2652)	-0.0558 (0.1578)	0.0142 (0.1627)	0.0275 (0.2115)
Type dummy	-	0.0487 (0.0840)	0.0533 (0.0851)	0.0431 (0.0888)
Time dummies	[0.000]	[0.000]	[0.000]	[0.000]
Wald (χ^2)	1.049 [0.592]	16.95 [0.031]	15.84 [0.045]	12.80 [0.119]
Sargan test	Not reliable	Not reliable	Not reliable	Not reliable
AR(2) test	[0.588]	[0.542]	[0.602]	[0.940]
Instruments	Size(2,3);ΔSize(1,1)	Size(2,3);ΔSize(1,1) Roa(0,3);ΔRoa(0,0) Mix(0,3);ΔMix(0,0) Eff(0,3); ΔEff(0,0)	Size(2,3);ΔSize(1,1) Roa(1,3);ΔRoa(0,0) Mix(1,3);ΔMix(0,0) Eff(1,3); ΔEff(0,0)	Size(2,3);ΔSize(1,1) Roa(2,3);ΔRoa(1,1) Mix(2,3);ΔMix(1,1) Eff(2,3); ΔEff(1,1)

Notes: ***,**,* denotes coefficient significantly different from zero at the 1, 5 and 10% level, respectively. Variables are transformed using first differencing. Results reported are estimated using robust standard errors (in parentheses). Figures in square brackets denote p-values. Wald is a test of joint significance of the independent variables (except constant and dummies). AR(2) is a N(0,1) test of no serial correlation in the residuals of the differenced model. Sargan is a test of overidentifying restrictions, asymptotically distributed as a chi-square under the null of instrument validity. The number of overidentifying restrictions is given in parentheses. As instruments, lagged levels of the variables are used (consistent with moment conditions [3.6]), together with lagged differences of the variables, consistent with [3.11]. No. of firms= 79; No. of observations = 395 (balanced panel).

Table A2: Two-step GMM- SYS estimation results
(forward orthogonal deviations transformation)
Dependent variable= ***Grw*** $_{i,t}$

	Univariate	**Multivariate**		
	(9)	**(10)**	**(11)**	**(12)**
Regressors		Exogenous	Predetermined	Endogenous
$Size_{t-1}$	-0.0221 (0.0377)	-0.0033 (0.0187)	-0.0084 (0.0196)	-0.0034 (0.0201)
Grw_{t-1}	0.0013 (0.0148)	-0.0199 (0.0182)	-0.0146 (0.0175)	-0.0102 (0.0205)
Roa_{t}	-	-1.382 (1.245)	-2.585 (1.804)	-3.412 (5.326)
Roa_{t-1}	-	0.4444 (1.669)	0.2194 (2.001)	0.9755 (3.373)
Mix_{t}	-	0.6059* (0.3449)	0.7731* (0.4093)	0.5311 (0.3787)
Mix_{t-1}	-	-0.3783 (0.3763)	-0.3861 (0.3827)	-0.3640 (0.3016)
Eff_{t}	-	0.3929** (0.1971)	0.3430** (0.1756)	0.4780 (0.4041)
Eff_{t-1}	-	-0.2592 (0.1999)	-0.3189 (0.2118)	-0.4364 (0.3012)
Constant	0.1204 (0.2369)	-0.0784 (0.1503)	0.0104 (0.1693)	-0.0056 (0.2321)
Type dummy	-	0.0514 (0.0795)	0.0605 (0.0837)	0.0443 (0.0898)
Time dummies	[0.000]	[0.000]	[0.000]	[0.000]
Wald (χ^2)	0.3862 [0.824]	15.04 [0.058]	15.81 [0.045]	9.151 [0.330]
Sargan test	[0.002] (14)	[0.849] (86)	[0.408] (71)	[0.206] (56)
AR(2) test	[0.682]	[0.516]	[0.616]	[0.832]
Instruments	Size(2,3);ΔSize(1,1)	Size(2,3);ΔSize(1,1) Roa(0,3);ΔRoa(0,0) Mix(0,3);ΔMix(0,0) Eff(0,3); ΔEff(0,0)	Size(2,3);ΔSize(1,1) Roa(1,3);ΔRoa(0,0) Mix(1,3);ΔMix(0,0) Eff(1,3); ΔEff(0,0)	Size(2,3);ΔSize(1,1) Roa(2,3);ΔRoa(1,1) Mix(2,3);ΔMix(1,1) Eff(2,3); ΔEff(1,1)

Notes: ***,**,* denotes coefficient significantly different from zero at the 1, 5 and 10% level, respectively. Variables are transformed using forward orthogonal deviations, as suggested by Arellano and Bover (1995). Results reported are estimated using robust standard errors (in parentheses). Figures in square brackets denote p-values. Wald is a test of joint significance of the independent variables (except constant and dummies). Sargan is a test of overidentifying restrictions, asymptotically distributed as a chi-square under the null of instrument validity. The number of overidentifying restrictions is given in parentheses. AR(2) is a N(0,1) test of no serial correlation in the residuals of the differenced model. As instruments, lagged levels of the variables are used (consistent with moment conditions [3.6]), together with lagged differences of the variables, consistent with [3.11]. No. of firms= 79; No. of observations = 395 (balanced panel).

Table A3: GMM- DIF estimation results

Dependent variable= $Grw_{i,t}$

	Univariate (13)	Multivariate (14)	(15)	(16)
Regressors		Exogenous	Predetermined	Endogenous
$Size_{t-1}$	-1.043** (0.4803)	-0.8763*** (0.1020)	-0.9391*** (0.0683)	-0.9756*** (0.1267)
Grw_{t-1}	-0.0495 (0.0380)	0.0136 (0.0403)	0.0037 (0.0352)	-0.0005 (0.0323)
Roa_t	-	-3.045** (1.229)	-3.878 (2.557)	-6.852** (2.814)
Roa_{t-1}	-	-2.818** (1.411)	-2.816*** (1.069)	-1.809 (2.242)
Mix_t	-	0.6866** (0.3273)	0.6181 (0.4409)	0.0418 (0.8647)
Mix_{t-1}	-	0.1931 (0.2403)	0.2555 (0.2475)	0.2350 (0.7435)
Eff_t	-	0.2342 (0.2316)	0.1777 (0.2511)	0.0880 (0.3114)
Eff_{t-1}	-	-0.0932 (0.0864)	-0.0977 (0.1253)	-0.2204 (0.2758)
Constant	0.1913*** (0.0265)	0.1467*** (0.0227)	0.1453*** 0.0406	0.1418** (0.0698)
Type dummy	-	0.0409** (0.0182)	0.0421** (0.0205)	0.0452 (0.0314)
Time dummies	[0.000]	[0.000]	[0.000]	[0.000]
Wald (χ^2)	25.15 [0.000]	180.4 [0.000]	299.2 [0.000]	123.2 [0.000]
Sargan test	[0.000] (8)	[0.225] (62)	[0.099] (47)	[0.046] (32)
Sargan stat	29.87	70.08	59.87	46.62
AR(2) test	[0.049]	[0.704]	[0.479]	[0.567]
Instruments	Size(2,3)	Size(2,3); Roa(0,3) Mix(0,3); Eff(0,3)	Size(2,3); Roa(1,3) Mix(1,3); Eff(1,3)	Size(2,3); Roa(2,3) Mix(2,3); Eff(2,3)

Notes: ***,**,* denotes coefficient significantly different from zero at the 1, 5 and 10% level, respectively. Results reported are estimated using robust standard errors (in parentheses). Figures in square brackets denote p-values. Wald is a test of joint significance of the independent variables (except constant and dummies). Sargan is a test of overidentifying restrictions, asymptotically distributed as a chi-square under the null of instrument validity. The number of overidentifying restrictions is given in parentheses. AR(2) is a N(0,1) test of no serial correlation in the residuals of the differenced model. As instruments, lagged levels of the variables are used (consistent with moment conditions [3.6]). No. of firms= 79; No. of observations = 395 (balanced panel).

Table A4: Pooled OLS and fixed effects estimation results

Dependent variable= $\boldsymbol{Grw}_{i,t}$

	Pooled OLS (17)		Fixed effects (18)	
Regressors	Univariate	Multivariate	Univariate	Multivariate
Sze_{t-1}	0.0022 (0.0038)	0.0007 (0.0044)	-0.3841*** (0.0392)	-0.4072 (0.0288)
Grw_{t-1}	-0.0119 (0.0186)	-0.0060 (0.0172)	-0.0194 (0.0168)	-0.0217 (0.0197)
Roa_{t-1}	-	-0.7125 (0.9520)	-	-1.875* (1.070)
Roa_{t-1}	-	1.647 (1.241)	-	-1.036 (1.551)
Mix_t	-	0.4001 (0.3437)	-	0.7212** (0.3484)
Mix_{t-1}	-	-0.4151 (0.3709)	-	-0.1135 (0.3090)
Eff_t	-	0.2988* (0.1586)	-	0.2554 (0.2295)
Eff_{t-1}	-	-0.2238 (0.1845)	-	-0.2148* (0.1252)
Type dummy	-	0.0513 (0.0285)	-	-
Time dummies	[0.000]	[0.000]	[0.000]	[0.000]
Wald (χ^2)	0.6156 [0.735]	17.48 [0.026]	101.8 [0.000]	441.4 [0.000]
AR(2) test	[0.486]	[0.552]	[0.020]	[0.021]

Notes: ***,**,* denotes coefficient significantly different from zero at the 1, 5 and 10% level, respectively. The t-ratios are based on robust standard errors (in parentheses); p-values are in square brackets. Wald is a test of joint significance of the independent variables (except time dummies). AR(2) is a N(0,1) test of no serial correlation in the residuals of the untransformed model. No. of firms= 79; No. of observations = 474 (balanced panel).

Table A5: Sample banks

Company name
BERGSLAGENS SPARBANK AB (PRIVAT)
FÖRENINGSSPARBANKEN SJUHÄRAD AB
FÖRENINGSSPARBANKEN SÖDERHAMN AB
IKANOBANKEN AB
NORDEA BANK SVERIGE AB (PUBL)
SKANDIABANKEN AB
SEB AB
SPARBANKEN GRIPEN AB
SPARBANKEN LIDKÖPING AB
SPARBANKEN SKARABORG AB
SWEDBANK AB
SVENSKA HANDELSBANKEN AB
TJUSTBYGDENS SPARBANK AB
VARBERGS SPARBANK AB
VIMMERBY SPARBANK AB
ALMUNDSRYDS SPARBANK
ALSKOGS SPARBANK
ATTMARS SPARBANK
BJURSÅS SPARBANK
BURS PASTORATS SPARBANK
DALHEMS SPARBANK
EKEBY SPARBANK
ESKELHEMS SPARBANK
FALKENBERGS SPARBANK
FARSTORPS SPARBANK
FRYKSDALENS SPARBANK
GARDA-LAU SPARBANK
GLIMÅKRA SPARBANK
GÖTERYDS SPARBANK
HISHULTS SPARBANK
HUDIKSVALLS SPARBANK
HÄRADSSPARBANKEN MÖNSTERÅS
HÖGSBY SPARBANK
IVETOFTA SPARBANK I BROMÖLLA
JÄRVSÖ SPARBANK
KINDA SPARBANK
KRISTIANSTADS SPARBANK
KYRKHULTS SPARBANK
LAHOLMS SPARBANK
LEKEBERGS SPARBANK
LEKSANDS SPARBANK
LÅNGASJÖ SOCKENS SPARBANK
LÖNNEBERGA SPARBANK
MARKARYDS SPARBANK
MJÖBÄCKS SPARBANK
NORDALS HÄRADS SPARBANK
NORRBÄRKE SPARBANK
NÄRS SPARBANK
RÖKE SOCKENS SPARBANK
SALA SPARBANK
SIDENSJÖ SPARBANK
SKATELÖVS OCH VÄSTRA TORSÅS SPARB.
SKÅNES FAGERHULTS SPARBANK
SNAPPHANEBYGDENS SPARBANK
SPARBANKEN I KARLSHAMN
SPARBANKEN NORD
SPARBANKEN SYD
SPARB. SÖRMLAND SÖRMLANDSBANKEN
SPARBANKEN TANUM
SPARBANKEN TRANEMO
SPARBANKEN VÄSTRA MÄLARDALEN
SÖDRA DALARNAS SPARBANK
SÖDRA HESTRA SPARBANK
SÖLVESBORG MJÄLLBY SPARBANK
TIDAHOLMS SPARBANK
TUNA-VENA SPARBANK
TYRINGE SPARBANK
ULRICEHAMNS SPARBANK
VADSTENA SPARBANK
VALDEMARSVIKS SPARBANK
VALLBY SPARBANK
WESTRA WERMLANDS SPARBANK
VINSLÖVS SPARBANK
VIRSERUMS SPARBANK
ÅLEMS SPARBANK
ÅRYDS SPARBANK
ÅSE OCH VISTE HÄRADS SPARBANK
ÅTVIDABERGS SPARBANK
ÄLMEBODA SPARBANK