Analysis of Regional Financial Institutions in the Environment of Structural Changes

戴, 建中

https://doi.org/10.15017/1500490

出版情報：九州大学, 2014, 博士（経済学）, 課程博士
バージョン：

Ph.D. Dissertation

Analysis of Regional Financial Institutions in the Environment of Structural Changes

Dai Jianzhong

Graduate School of Economics, Kyushu University

Acknowledgement

At first, I would like to express my gratitude to my respectable supervisors Professor Masao Nakata and Professor Akinori Isogai. Professor Nakata supervises me from the beginning of my PhD study. He guided my research from the data collection to the writing of the dissertation. Even after he left Kyushu University to Seiko University; he continues to guide my PhD study. Very often he traveled between Tokyo and Fukuoka to give me seminars. Before each conference of my PhD defending, he gave me detailed instructions about the presentation. Professor Isogai took the responsibility of supervising my PhD study after Professor Nakata left Kyushu University. He gave me many valuable instructions about the dissertation. After I returned to my country, he kindly takes care with the necessary administrative procedures so that my return to Kyushu University to defend my dissertation is possible.

Second, I would like to thank Professor Yasuhiro Horie. He is one of my supervisors before he retired from Kyushu University. He helped me to determine the topic of this dissertation and taught me many things about the regional financial institutions in Japan.

I also would like to express my gratitude to Professor Kunio Urakawa and Nobuaki Hori, Department of Economic Engineering, Faculty of Economics in Kyushu University. They kindly agreed to be the members of my dissertation advisory committee. Their critical comments helped me to improve the quality of the dissertation

Parts of the dissertation are presented or published by referred journals. I also express my appreciation here to all the commentators and referees for their valuable opinions on the corresponding parts of the dissertation.

During my study in Kyushu University, I was impressed by the favorable environment of research and the instructiveness of the classes I have taken. Without these helps, the completion of this dissertation is impossible.

Finally, I also wish to thank Kyushu University and her alumni association for the financial supports they gave to me. These supports greatly reduced my financial burden of PhD study.

Dai Jianzhong

2014, 12

Abstract

Regional banks play a very important role in the economy. Thus, many countries have implemented various programs to restructure regional banks in their attempt to reform their economies. This dissertation is an attempt at exploring the effect of these financial restructuring programs on regional banks. The dissertation focuses on the effects of financial restructuring on three important areas of regional banks: credit supply, risk and the efficiency of the banks themselves. Compared with the previous studies, the dissertation touched upon more areas connected with the healthy operation of a banking system. The dissertation also utilizes various updated statistical and econometric techniques and thus is a more thorough and robust analysis of the topic.

The dissertation uses Shinkin banks in Japan and city commercial banks in China as objects of the research. The first three chapters of the dissertation analyze the measurement and determination of efficiency and productivity changes in Shinkin banks in Japan during the period of 2001 to 2008 (Fiscal year). The dissertation attempts to determine how the efficiency and productivity of Shinkin banks changed during the period and how these changes are related to financial restructuring.

The dissertation uses a two-stage approach to achieve this aim. At the first stage, which is based on the theoretical background explained in Chapter 3, Chapter 4 offers a robust estimation of the efficiency and productivity changes in Japan's Shinkin banks from 2001 to 2008 (Fiscal year). Hyperbolic-oriented distance measurement is used to measure the efficiency. A non-parametric method called Data Envelopment Analysis (DEA) is used to estimate the distances. Based on the estimated efficiency scores, the Malmquist index, which is a widely used index, is employed to measure the productivity changes in Shinkin banks. To overcome the shortcomings of the DEA method, a non-parametric, smooth bootstrapping method is utilized to establish the confidence intervals and significance levels for the estimated results. The efficiency scores and Malmquist index are decomposed to examine the sources of the trends in efficiency and productivity changes. The scores and the index are also grouped

according to asset levels and market powers of sample banks to explore the influence of bank scale and market power of Shinkin banks on their efficiency and productivity changes, respectively.

The dissertation identifies the trends in efficiency and productivity changes in Japan's Shinkin banks during the study period by analyzing the estimated scores. The dissertation shows that on average, the efficiency under the assumption of variable return to scale (*vrs*) significantly declined from 2005 to 2008 (Fiscal year). However, the scale economy (sc) significantly increased in the same period and overcame the negative influence of the *vrs*, thereby allowing the efficiency measured under the assumption of constant to scale (*crs*) to increase significantly. The dissertation also determines that in the latter half of the study period, productivity significantly declined, primarily because of the deterioration of technical efficiency, whereas scale efficiency significantly improved. Grouping the total sample according to the levels of assets and competition reveals more details of the trends of efficiency and productivity changes in Shinkin banks.

Based on the estimated scores obtained in Chapter 4, Chapter 5 further explores the effects of the merger and acquisition (M&A) activities that occurred in Japan during the period of 2001–2004 (Fiscal year)on the efficiency and productivity changes in Shinkin banks during the period of 2005–2008 (Fiscal year). The efficiency scores estimated in Chapter 4 are initially used as dependent variables. A truncation model is used to estimate the parameters. Subsequently, the Malmquist index and its components estimated in Chapter 4 are used as dependent variables. The dissertation utilizes a system of equation approach to analyze the determination of productivity changes. In the analysis, to cope with the non-parametric approach used in estimating the dependent variables and to deal with the complex problems involved in the model, the dissertation uses a semi-parametric bootstrapping approach to test the significance of the coefficients in the model.

The dissertation shows that M&A incidents have no significant effects on the Malmquist index. However, they have significant effects on efficiency scores and two components (technical and scale efficiency) of the Malmquist index. The dissertation

also discovers that several other factors related to M&A have significant positive effects on the efficiency and productivity changes in Shinkin banks. For example, the dissertation finds that the medium sized banks are ranked first in scale economy and there is a hump-shaped relationship between bank assets and efficiency. This empirical result implies that the policy maker should try to find the optimal scale for a regional bank and only encourage M&A among small banks

The last two chapters are concerned with how the implementation of the capital adequacy ratio (CAR)—an important bank risk control policy—affected the loan supply ability and safety of city commercial banks in China. Chapter 6 surveys the theoretical literature on the endogenous relationship between loan supply, bank capital, non-performing loans (NPLs), and economic environments. Based on this literature review, the dissertation attempts to build a theoretical model to describe this endogenous relationship.

Chapter 7 checks the effects of capital regulation on the loan supplies empirically by using city commercial banks in China during the period of 2005–2008 (Calendar years) as samples. The dissertation develops a simultaneous equation model (SEM) to analyze the endogenous relationship between capital changes, loan growth, and NPLs. The dissertation shows that during the sample years, capital condition became a strict constraint for loan growth of these banks. This constraint also became increasingly strict and significant across the sample years as the implementation of the policy continued.

In conclusion, the dissertation supports the view that restructuring programs are generally beneficial to the improvement of productivity and soundness of regional banking systems. This conclusion may offer some theoretical support for further bank restructuring in these two countries, as well as in other countries.

Key words: productivity, banking, DEA, smoothed bootstrapping, Japan, capital adequacy ratio; loan supply; non-performing loan; simultaneous equation model;

JEL codes: C14; D24; G21; C31; E51; G21;

Contents

List of Table

Lists of Figures

Chapter 1 Introduction

1.1 Research background

This dissertation focuses on the analysis of the effects of financial restructuring on regional banks.

In almost all economies, Small and medium enterprises (SMEs)[1] play a very important role in almost all economies. Although SMEs are relatively less technically advanced and contribute less to the economy per company as compared to large corporations, as a whole, they account for a larger contribution to the economy. For example, in China, from 2001 to 2005, the added value of SMEs increased by 28% on average annually, whereas the average annual growth rate of GDP was 9.5%. By the end of 2008, SMEs accounted for 99.6% of the number of total enterprises, about 60% of the GDP, and 50% of the tax. SMEs also accounted for 68.85% of the total import and export values and about 80% of overseas investment (Shen, Shen et al. 2009). According to the 2012 version of "The white paper of SMEs" published by the Japanese government, Japan has 4.198 million SMEs, accounting for 99.7% of the total number of enterprises. They have 27.84 million employees and comprise about 70% of the total employment in Japan. In manufacturing industry, they contribute 4.84 billion Yen of the added value, about 50% of the total added value.

More importantly, in modern economies, large corporations heavily rely on

[1]In China, SMEs are defined as enterprises with employees less than 2000, or sales less than RMB￥30000, or total assets less than RMB￥ 40000. In Japan, according to the "Basic law of SMEs", SMEs are enterprises with capital less than 50 million Yen for retail and service industry, 100 million Yen for wholesale industry, and 300 million for manufacturing, construction, and transportation industry. They have less than 50 employees for the retail industry, 100 for the wholesale and service industry, and 300 for the manufacturing industry.

SMEs for intermediate inputs and outsourcing. With the support of SMEs, large corporations have substantially reduced their costs and raised their competitive power. Japan (after World War II) and China (since 1990s) could not have achieved such success in manufacturing exports without the support of a relatively comprehensive SME system in their domestic economy.

Unlike large corporations, SMEs are usually labor intensive. Thus, they have a disproportionally larger share of contribution to the employment than their share of corporate assets. In China, they account for 75% of the urban employment, and their share of GDP is 60% in 2008. In Japan, they account for about 70% of the total employment, but only account for 50% of the added value in the manufacturing industry.

SMEs are especially crucial for the economies of the regions of their location. Unlike large corporations whose businesses are nationwide, SMEs usually operate over limited geographical areas. Therefore, their business strategies respond quicker to the requirements of regional economies.

As shown in the United States and many other countries, many SMEs are innovative and high technological firms. Changing operational directions is relatively easy for SMEs. A number of famous large corporations began as SMEs, especially in the information and other high technology industries. Therefore, SMEs are essential to the vitality and technological advances in an economy.

Supplying financial sources to SMEs is critical because of their importance in the economy. However, in many countries, SMEs experienced difficulties in obtaining financial support. This has become a severe barrier for the development of SMEs. For example, in China in 2006, no more than 0.5 million SMEs were able to access bank loans. This figure means over 98% of the total 40 million SMEs in China were not able to obtain bank loans (Lin 2007). In Japan, SME's access to finance is also limited. They accounted for about 70% of the balance of credit to enterprises in 2013 (see Table 1.1), although they accounted for 99.7% of the total number of enterprises. After financial liberalization, the sources of financing of large corporations in Japan widened drastically, but for SMEs, bank loans still accounted for 70% of their financial needs. The reason for the difficulty of SMEs in obtaining financial support is because traditional lending techniques used by large banks are not suitable for the needs of SMEs.

Table 1.1 Balance of Credit to SMEs

Unit: 1000 billion Yen

	2009	2010	2011	2012	2013
Credit to SMEs	219.711	216.266	213.799	212.869	216.058
Shinkin Banks	42.093	41.455	41.313	40.920	41.204
Equipment finance	16.974	16.824	16.865	17.103	17.633
Financing for working capital	25.118	24.631	24.448	23.817	23.571
Domestic banks	177.619	174.811	172.486	171.949	174.854
Equipment finance	80.179	79.487	78.833	80.071	82.464
Financing for working capital	139.532	136.779	134.966	132.798	133.594
Total credit to enterprises	317.665	311.487	311.511	314.898	320.150
Equipment finance	93.901	93.176	92.611	94.193	96.349
Working capital Financing	223.765	218.312	218.900	220.705	223.801

Source: 2013 annual overview of the Central Shinkin banks

For a long time after the World War II, a number of countries encountered serious malfunctions in their banking systems because the government interfered heavily in the financial industry. These malfunctions caused severe problems in the economies of these countries. The results were low efficiencies of financial institutions, high cost of financing for enterprises and consumers, and limited sources of financing, especially for SMEs. Since the latter part of the 20^{th} century, an increasing number of governments worldwide have become aware of the problems in their banking systems and have correspondingly implemented various programs to restructure the system. In the 1980–90s, the focus was placed on the liberalization of the banking system. In many countries, the separation of operations of different kinds of financial institutions was eliminated. Interest rate ceilings were waived. Government interferences in loan decisions of banks were cancelled. These policies were designed with the idea of fostering significant improvements in the profitability of projects that banks invested in, thereby increasing the efficiency of the banks themselves, as well as the efficiency of the economy as a whole. In these processes, the governments were especially interested in improving the financial environment for SMEs. In addition to directly offering financial support to SMEs, governments also offered various kinds of support for the development of regional financial institutions.

However, many severe banking crises occurred in the 1900s and 2000s, including the Latin American debt crisis occurred in the early 1980s, the Mexican financial crisis in 1994, and the south eastern Asia crisis in 1998. There were also national financial crises with huge international impact. The most serious ones were “credit crunch” in the United States in 1992–1993, and a similar but more severe and lasting crisis in Japan in the latter part of the 1990s. Upon entering the 21st century, a much more severe and wide-spread international crisis broke out in 2008, which originated from the prime debt crisis in the United States. In these crises, SMEs usually suffered the most because, as many researchers have found, the financing for these enterprises was usually the first to be withdrawn and reduced by banks. Many experts suspected that the financial liberalization carried out in the latter part of the 20th century was partly responsible for these crises. Thus, since the late 1980s, governments in many countries and international supervisory organizations have been trying to implement some forms of risk regulations on banks. The most important one is the risk weighted capital adequacy ratio (CAR) regulation proposed by the Basel Committee on Banking Supervision in 1988. Since its implementation, the CAR regulation has been adopted by an increasing number of countries worldwide. This policy requires banks to maintain their capital/risk adjusted assets ratio always above the minimum level. Proponents of this policy hoped it would force banks to be more willing to control their risks.

It is quite interesting to check whether these restructuring programs have influenced the function of regional banks. The dissertation uses Shinkin banks in Japan and city commercial banks in China as objects of the research. Both are typical regional financial institutions in each country. At first glimpse, these two countries are entirely different. One is a developed country, while the other is a developing country in rapid transformation. However, after entering the 21st century, both countries experienced drastic financial restructuring. Before the financial restructuring, despite the fundamental differences in their economic systems, the governments in both countries had strict control over financial institutions. This control leads to significant inefficiency and the accumulation of a huge amount of NPLs in the financial systems in both countries. Coinciding with the introduction of CAR in the 1990s, the problems in the banking systems became a serious threat to the health of economies of both countries. In Japan, in the latter part of the 1990s, Japanese banks were severely under-capitalized because of their deletion of NPLs. This weakened their ability to supply loans and caused a “credit crunch.” In China, although the situation was not as dire, huge NPLs also depleted the capital of the banks and constrained their ability to supply loans. At the beginning of the 21st century, the governments of both countries realized that the old financial systems required fundamental changes. In Japan, after the economic bubble burst, the government launched a series of financial liberalization programs. In China, the government implemented important reform programs that would gradually transform the old government-directed financial system into a more market-oriented financial system.

Research in this field will not only enrich existing literature on banking, but also illuminate the design of financial restructuring programs in the future, because the more we understand the results of the already implemented program, the more appropriate the restructuring programs proposed in the future will be. These types of studies can help us understand how bank restructuring affects the efficiency of banks, whether the CAR regulation is effective in controlling the risk-taking of banks and how it is accomplished. Based on these understanding, we can further understand how the CAR regulation will influence the effects of the monetary policy (the so-called credit channel). We may also use our understanding of the effects of capital regulation to design anti-business cycle macroeconomic policies by adjusting the toughness of the capital regulation. For SMEs, understanding the effects of financial restructuring on regional banks will let us better understand the effects of financial restructuring on SMEs. This will help us design financial restructuring programs more favorable to SMEs.

1.2 Theme and methodology of the dissertation

The dissertation analyzed the effects of financial restructuring on regional banks using Shinkin banks in Japan and city commercial banks in China as the samples.

The performance of a banking system can be evaluated from three sides. It can be evaluated by how much credit a banking system has supplied to the economy. A well-functioning banking system should be able to satisfy all kinds of financing needs in the economy, whether for large corporations, SMEs, or consumers. However, when banks themselves are in trouble (especially when they have accumulated a huge volume of NPLs and are in shortage of capitals), their credit supply ability are jeopardized. This so called "credit crunch" phenomenon first caught the attention of economists in the United States in the early 1990s. Japan (after 1997) and China (in the 2000s) also faced the same problem for a long time. Even if there is no shortage of credit supplies in general, there may be structural insufficiency of credits. Some customers (especially SMEs) find it difficult to obtain credit from banks. During the "credit crunch," they suffer most severely from credit reductions.

The second standard is how much risk a banking system takes. Banks that take too much risk will inevitably end in heavy losses of loans. This will not only waste many of the society's resources, but may also weaken the ability of banks to supply credit and may cause serious problems in the economy. However, this standard sometimes is contradictory to the first one. Most of the credits supplied by banks to the economy are high-risk assets; thus, the more loans supplied to the society, the more risk the banks take. A society has to balance these two aims.

The third standard of measurement concerns the efficiency of the banking system itself. Obviously, any well-functioning industry should operate in a way that uses the least social resources to satisfy the demand for their products/services from the society. Furthermore, efficiency will strengthen the safety of the banks; therefore, it will help banks to achieve the first two aims. An efficient bank will have more profits to replenish capitals; thus, it will be safer and have a stronger ability to apply credits.

Even non-profit seeking financial institutions such as Shinkin banks require the maintenance of a certain level of efficiency. Otherwise, they cannot maintain their operation. On the other hand, safe banks will enjoy a good reputation in the capital market and will be rewarded generously. They will be able to borrow more from the capital market and pay less risk premium, which will allow them to be more efficient. However, if they do not meet the minimum requirement of safety (such as the minimum CAR level), they will be punished severely by the government and have to pay high risk premium when they borrow money from the capital market.

Based on this division of bank functions and relationship among them, the dissertation focuses on the effects of financial restructuring on three important areas of regional banks: credit supply, risk, and efficiency of the banks themselves. After a concise description of the theoretical foundation of the dissertation, the following three chapters analyze the measurement and determination of efficiency and productivity changes using as sample the Shinkin banks in Japan during the period of FY 2001–2008. The dissertation attempts to determine how the efficiency and productivity changes of Shinkin banks changed during the period and how these changes were related to financial restructuring in the same time. The last two chapters are concerned with how the implementation of CAR—an important bank risk control policy—have affected the loan supply ability and safety of city commercial banks in China during the period of 2005–2008.

The reasoning behind the sample selection is that the major problems in the regional banking systems of Japan and China now are different despite all having experienced bank restructurings since the 1980s. As an advanced country, Japan has a relatively mature financial system. Although Japanese banks were very radical in their operation during the 1980s and early 1990s and accumulated huge volumes of NPLs, this problem was solved at the beginning of the 21st century. Currently, over-risk taking is no longer a big concern of Japanese supervisors. However, since the Japanese economy entered a stage of relatively stable but low growth rate, how to maintain and improve the productivity of regional financial institutions becomes an urgent problem. Thus, for Shinkin banks of Japan, the dissertation focuses on the evolution of their productivities in the 2000s.

On the contrary, from the 1980s, the Chinese economy has rapidly changed and transformed. The economic growth rate is very high but is accompanied by high fluctuations. This growth trend greatly increases the risk taken by the Chinese financial institutions. Furthermore, compared with their colleagues in Japan and other developed countries, they are more aggressive in their operations and more eager to offer loans to customers. Thus, determining the means to curb the impetus of banks to over-supply loans is the first concern of Chinese supervisors. Therefore, in the analysis of the city commercial banks in China, I focus on the relationship between bank loan supplies and bank conditions.

The dissertation uses a non-parametric approach to estimate the efficiency and productivity change of Shinkin banks in Japan. Compared with the more traditional econometric approach, non-parametric approach does not assume any particular form of function. This characteristic sometimes makes it superior to the econometric approach when it is difficult to fix a function form for the object of the analysis. However, this characteristic also has some shortcomings, such as being unable to accommodate error and random effects in the model and its results are very sensitive to outliers.

All databases used in this dissertation are panel data. However, unlike most of panel databases used in other research projects, the database used in this dissertation all have short time lengths (small T) but large number of individuals (large N). The time length used in the analysis of the efficiency and productivity changes of Japan's Shinkin banks is eight, but the number of sample banks in each year is about 200–300. The time length used in the analysis for the loan supply of the city commercial banks in China is four, but the number of sample banks in each year is about 80–120.

When we have a panel database which is large enough at least in one side (either T or N), we can do much more than simply using the one-way effect (either fixed or random) panel models that only assume varying intercept terms. We can use two-way effect (either fixed or random) panel models that allow not only the intercept term but also some of the coefficients to vary over time or across different individuals, or we can run a series of regressions for each year or each individual (group). Both approaches should produce similar results. However, the latter may reveal more interesting information. For example, if we stack the cross sectional regressions year by year, by examining the results for each year, we can check not only for the presence of noticeable changes in some coefficients, but also for changes in the significance (the *t* value) of these coefficients. We can also check the change of total fitness of the model (the R square) during the sample period. By observing these changes during the period when an important policy is implemented, we can deduce the effects of this policy on the behaviors of target agents.

Due to the characteristics of the database used in the dissertation, the dissertation runs a series of cross-sectional models to fulfill the purpose of the research. However, the dissertation does not use single cross-sectional equation approach to estimate the model. Rather, the dissertation uses a system of equations approach to estimate the model. Unlike the single equation model that only uses information in the equation, a system of equations approach can use the entire set of information in the system (correlation among different equations, connection between coefficients in different equation, endogenous relationship among some of the variables in the model, etc.). Therefore, the estimation of the model may be more efficient and robust. In the analysis on the efficiency and productivity change in Shinkin banks in Japan, the dissertation uses the connection between the equations for the total score of productivity change and its components. In the analysis of loan supply of city commercial banks in China, the fact that the residuals between different equations are correlated is explored.

For cross-sectional econometric models, the problems that may arise are heterogeneity, multi-linearity, errors in the variables, endogenous independent variables, etc.. In the econometric models for determining the efficiency and productivity changes in Shinkin banks in Japan, the problems are biased-dependent variables and endogenous-independent variables. For by construction, the efficiency and productivity change scores and their components are biased estimations of the real scores. These scores are also correlated with each other so that their explanatory variables are correlated with contemporary errors terms. Because the scores are estimated by non-parametric approach, the problems are too complex to be handled by ordinary econometric techniques. Therefore, the dissertation uses a semi-parametric bootstrapping approach to solve the problems. In analyzing the loan supply of city commercial banks in China, the problem is the endogenous relationship between loan supply and bank conditions (CAR and NPL ratio). The dissertation uses a SEM approach to deal with the problem.

1.3 Innovation and contribution of the dissertation

There is now a huge volume of researches on the financial institutions in Japan and China. Most previous studies focused only on one aspect of the function of financial institutions. The dissertation attempts to conduct a more thorough analysis of the effects of financial restructuring on a certain kind of financial institution. Compared with existing literature in this field, the dissertation touched upon more areas connected with the healthy operation of a banking system. The dissertation includes most of the functions of a financial institution: loan supply, risk, and efficiency of banks. This study also utilizes many updated statistical and econometric techniques. Hence, it is a more thorough and robust analysis of this topic. In some cases, using the more robust methods has assisted the dissertation to obtain results similar to other related studies. In these cases, the dissertation offers more robust evidence for the conclusions. In other cases, it enables the dissertation to obtain results quite different from that of the other research.

Some papers analyze the efficiency and productivity changes in Japanese financial institutions. Most of these studies focus on the estimation of the economy of scale and effect of M&A on productivity. Several of these studies have used the non-parametric DEA approaches. For example, Fukumaya (1993) analyzed the technical and scale economy of 143 Japanese banks in FY 1991. The sample includes city and regional banks. He showed that the major cause of inefficiency was pure efficiency, not diseconomy of scale. He also found a significantly positive correlation between bank scale and economy of scale. Fukuyama (1996) also analyzed the technical efficiency and return to scale for a sample of 435 credit unions in FY 1992. In his paper in 1993, he found substantial inefficiencies in credit unions in Japan. The major cause of inefficiency is pure inefficiency. Drake and Hall (2003) estimated the efficiency of Japanese banks in FY 1996 for a sample of 149 banks. The sample included all “ordinary banks” (city and regional banks) operating in Japan. He found that large city banks have very little room for improvement in pure efficiency or economy of scale. Thus, encouraging M&A among large banks will not significantly improve the efficiency of large banks. On the contrary, the smallest banks have the lowest level of economy of scale. Thus, they support M&A among small banks. Fukumaya and Weber (2008) continued the study of Fukuyama (1993) for the period of FY 2002–2004. Their sample includes both city and Shinkin banks. Their major interest was in the estimation of shadow prices of NPLs. They found that for financial institutions, the results of efficiency estimation would be notably changed if you include NPLs in the model. Drake et al. (2009) estimated the efficiency of Japanese financial institutions during the period of FY 1995–2002 and compared efficiency scores under different assumptions of banks. They found that the results of efficiency estimation would be quite different using different assumptions. Horie (2010) estimated the productivity changes in Shinkin banks of Japan during the period of FY 2005–2007and analyzed the relationship between productivity changes and operating areas. He found that productivity changes of Shinkin banks are significantly and negatively related to their operating areas.

However, all researchers except Horie’s (2010) focused on the efficiencies, not the productivity changes of Japanese financial institutions. Furthermore, those that used the DEA approach did not consider the randomness of estimates. We do not know the actual production function in the real world and thus, we can only estimate it from a selected sample. As in parametric estimates, different samples will lead to different estimates. These estimates will normally deviate from real values. However, the DEA approach is deterministic in nature and does not accommodate random factors in the estimation. In this dissertation, I try to use some recently developed bootstrapping techniques to obtain a more robust estimate of the efficiency and productivity changes and test the significance of the results.

To solve this problem, the dissertation uses the nonparametric bootstrapping approach suggested by Simar and Wilson (1999) to obtain robust estimates of the efficiency scores and the Malmquist index. Later in the analysis of the effects of M&A on the efficiency and productivity changes of Shinkin banks, the dissertation also uses semi-parametric bootstrapping approach by Simar and Wilson (2007) to deal with the complex relationship between the dependent variable and explanatory variables. To my knowledge, this is the first attempt of such kind in a study on the efficiency and productivity changes of Japanese financial institutions.

The second problem in existing literature is that when they use efficiency as the dependent variable to analyze the determination of efficiency, they only use OLS approach. It is well known that the estimated efficiency scores are truncated variables and thus, OLS is a biased estimator for them. To avoid this problem, this dissertation uses a truncation model to estimate the effects of M&A on efficiency scores of Shinkin banks. It also uses the Malmquist index as dependent variable to analyze the effects of M&A on productivity changes of Shinkin banks. The Malmquist index is not truncated.

The third innovation of the first part of this study is that I use the hyperbolical-oriented distance instead of input- or output-oriented Shephard distance as the measure of efficiency. This is also the first attempt to use this measurement for the case of Japanese financial institutions. Similar to the directional distance measurement used by Fukuyama and Weber (2008), hyperbolical-oriented distance considers both output and input efficiency. With this measure, we can avoid the problem of possible discrepancies between input- and output-oriented distances. Unlike the directional distance, it is also easy to be decomposed. In addition, hyperbolic-oriented distance is also closely related to the concept of profit, which is the conventional measure of efficiency.

There are also several researches about the relationship **o**f bank capital and loan supply in China. For example, Zhao and Wang (2007) used a cross-sectional model to analyze the effects of capital position on loan supply in China during the period of 1995–2003. Their sample included 12 Chinese banks and they found that capital positions had no significant effects on loan supplies.

Dai, Jin et al. (2008) focused on the analysis of the relationship between capital management, bank loan supply, and monetary policy in China during the period of 1998–2005. They used the four largest national banks ("The Big Four") as their sample and utilized a simultaneous model, which included both a supply and a demand equation for bank loan, in their analysis. Their results indicated that in China, the effect of monetary policy is asymmetric because of capital supervision.

Wu and Zhou (2006) analyzed the endogenous relationship between capital and risk taken by commercial banks by using a simultaneous model. Their sample included 14 large and medium banks in China between 1998 and 2004. Their study showed that the implementation of CAR regulation in China had negative significant effects on the risk taken by Chinese banks.

Wang and Wu (2012) also analyzed the effects of capital supervision on bank loans in China. They used unbalanced panel data of Chinese commercial banks during the period of 1998–2009 and established a reduced form panel data model. Their results showed that capital position was positively and significantly related with loan growth rate over the entire sample period. They also found that the effect of capital position on long growth was stronger for capital-constrained banks.

Xu and Chen (2009) built a dynamic stochastic general equilibrium (DSGE) model to simulate the dynamic relationship between bank credit and economic fluctuation in China using quarterly macroeconomic data from 1993 to 2005. The paper revealed that credit shocks were the major cause of fluctuations of short term consumption, loans, and real money balance.

This dissertation also develops a SEM composed of three equations. This approach offers a clearer description of the complex interrelationship between bank conditions, loan supply, and economic environment. The model used in this dissertation differs from the models mentioned above in that it directly considers the endogenous relationship between non-performing loans (NPL), loan growth, and capital changes. Loan growth will influence NPL growth and the accumulation of NPL will negatively affect bank's capital positions. The weakening of bank capital position will in return affect the bank loan growth.

I also make great effort to enlarge the sample. It is well known that the sample of Chinese banks with detailed data available is very small, especially before 2008. Instead of only relying on Almanacs or available database such as Bankscope, I used the annual reports of Chinese banks collected from the website. Compared with other papers related to the banking system in China, to my knowledge, the sample in this research is the largest.

For the analysis of the relationship between loan supply and bank conditions, most previous studies used time series, cross-sectional, or bank panel data combined with macroeconomic data analysis. In this research, I use a bank panel database combined with regional economic data. I believe the regional economic data are more suitable in analyzing regional financial institutions. Furthermore, to measure more precisely the economic environment that each regional bank faces, I use weighted average of regional economic data for banks that operate in more than one region. Furthermore, I do not use simple cross-sectional analysis or panel analysis. Instead I use stacked cross-sectional analysis. As explained above, this approach will reveal more about the effects of policy implementation on the analyzed objects.

1.4 The structure of the dissertation

The remainder of the dissertation is arranged in seven chapters.

Chapter 2 offers some background knowledge on the dissertation. To explain the importance of topics in this dissertation, I first describe the basic theories on the roles played by commercial banks, particularly regional banks. Second, I outlined the financial systems in Japan and China as well as their institutional evolution and policy environment changes over time. This outline will help those unfamiliar with financial systems in these two countries understand better the analysis in this dissertation.

Chapters 3 to 5 are concerned with the analysis of efficiency and productivity changes in Shinkin banks in Japan.

Chapter 3 describes the theories and techniques for non-parametric estimation of efficiency and productivity changes. The chapter also explains how to use the bootstrapping method to establish confidence intervals and significance levels for the estimates of efficiency and productivity changes obtained through non-parametric estimation approaches. This chapter will build theoretical and technical foundations for the productivity analysis of Shinkin banks in Japan.

Chapter 4 conducts a robust estimation of the efficiency and productivity changes in Japanese Shinkin banks from FY 2001 to FY 2008. A measurement called hyperbolic-oriented distance is used to measure efficiency. A non-parametric method called data envelopment approach (DEA) is used to estimate the distances. Based on the estimated efficiency scores, a widely used index called Malmquist index is employed to measure the productivity changes in Shinkin banks. The shortcoming of DEA mentioned previously is overcome by using a non-parametric, smooth bootstrapping method to establish the confidence intervals and significance levels for the estimated efficiency scores and Malmquist index. The efficiency scores and Malmquist index are decomposed to two and three components respectively to examine their sources. The estimated efficiency score is further grouped according to bank asset to explore the influence of bank asset on their efficiency. The Malmquist index is grouped according to market powers of Shinkin banks to explore the influence of market powers on their productivity changes.

Chapter 5 explores the effects of M&A that occurred in Japan during the period of FY 2001 to FY 2004 on the efficiency and productivity changes in Shinkin banks during the period of FY 2005 to FY 2008. A system of equation approach is utilized to analyze the problem. In the analysis, the efficiency scores and Malmquist index and its components, which were estimated in Chapter 4, are used as dependent variables. To deal with the characteristics of the estimated scores of efficiency, a truncation model is used to analyze the determination of efficiency. To cope with the non-parametric approach used in estimating the index as well as to deal with the complex problems involved in the model (measure errors and serial correlations in dependent variables and endogenous independent variables, etc.), I use a semi-parametric bootstrapping approach to test the significance of coefficients in the model.

Chapters 6 to 7 examine the effects of capital supervision on bank loan supply.

Chapter 6 surveys the theoretical literature on the endogenous relationship between loan supply, bank capital, NPLs, and economic environments. Based on this literature review, the dissertation builds a theoretical model to describe this endogenous relationship.

Chapter 7 provides an empirical check of the effects of capital regulation on loan supplies by using a sample of city commercial banks in China during the period of FY 2005 to FY 2008. An SEM based on the theoretical models built in Chapter 6 is used.

Chapter 8 presents the conclusions of this dissertation. Based on the conclusions, the dissertation provides several policy proposals for financial restructuring programs in the future. Finally, the weaknesses of the dissertation are identified and recommendations for future studies are provided.

Chapter 2 Role of regional financial institutions and its revolution

In this chapter, the necessary theoretical foundation for the importance of the aim of this dissertation is provided and background knowledge on the financial systems in Japan and China and their systematic evolution is offered. This foundation may help those unfamiliar with the financial systems in these countries better understand the analysis in this dissertation.

2.1 Role of commercial banks in an economy

Commercial banks play a key role in an economy. As a financial intermediary, banks mainly absorb funds from savers in the form of deposit and transfer them to investors in the form of loans. Contrary to other financial institutions, most fund suppliers of banks (depositors) are ordinary inhabitants who have no profound knowledge on investment theories and techniques or ability to collect information. Thus, banks monitor debtors in delegation of depositors and bear the risk of investment projects in many cases.

Apart from this major function, banks are the main provider of payment instruments and fund transferring services, foreign exchange transactions, and risk arbitrage contracts.

The importance of banking system has been widely acknowledged in real economy. However, in traditional microeconomic theories based on perfect information such as the Arrow–Debreu general equilibrium model, banks have no role to play. It was not until the innovation of imperfect information economics did some analysts begin to address the question why there is a necessity for the existence of banks. Leland and Pyle (1977) were the first to explain the necessity of banks by using imperfect information theories. Diamond (1984) established the foundation of the rationality of banks by using a delegation theory.

In an imperfect information world, banks can exist because they have a few information cost advantages compared with direct contracting between creditors and debtors. This is due to the fact that direct financing may have duplicating efforts in ex ante screening (Leland and Pyle 1977) or in ex post monitoring (Diamond 1984) by different creditors. An information asymmetry exists between creditors and debtors. Debtors normally have some private information on ex ante expected returns or ex post realized returns of invested projects. Creditors have to spend some information costs to screen investment projects or monitor the behaviors of debtors. When many debtors lend in a single project, duplicate information cost may be spent in this project. If a bank delegates all creditors to screen or monitor the project, then this duplication of information cost can be avoided. However, in this case, creditors have to spend a certain amount to screen or monitor the banks or banks have to spend a certain amount to signal the quality of their investment. Diamond (1984) proved that under certain conditions, the monitoring cost on banks by depositors or the signaling cost by banks is less than the cost saving advantages because of the delegation of banks; hence, it is beneficial to the creditors and society in general.

By solving the asymmetric information problem, banks can improve efficiency of distributing financial resources. Consequently, banks can affect economic growth in a country. The correlation between economic growth and banking system development is evident from statistical data. However, determining the causality relationship between these two variables econometrically is difficult. Whether economic growth motivates the growth of a banking system (demand following) or the improvement in banking system stimulates economic growth (supply leading) remains unclear. However, Patrick (1966) argued that in the early stage of economic development, it is mainly banking system development that leads economic growth. When the economy is fully developed, it is mainly the demand from the economy that causes the development of a banking system.

How well a banking system functions can be measured by three standards. The first standard measures how much credit a banking system has supplied to the economy. A well-functioning banking system should be able to satisfy the financing needs in the economy. Banks fulfill this aim by offering attractive deposit instruments to inhabitants. Researchers differ on how a banking system will affect savings. Some of them maintained that a highly advanced banking system may induce inhabitants to consume more and save less (such as by providing consumption credit). For example, Diamond and Dybvig (1983) established a three-period model for savings and showed that because of the liquidity offered by banks, the demand for savings for the purpose of unexpected expenditure in the later periods (precautionary saving) may decline. However, it is widely accepted that in the early stage of financial development, advances in the banking system will stimulate savings by offering more attractive financial instruments to inhabitants and making contact with banks easier. The financial regression theory proposed by Makinnon (1973) and Shaw (1973) pointed out that in a few developing countries, governments have deliberately kept the interest rate lower than the free market rate to offer implicit subsidy to manufacturers. This greatly hinders the development of the financial system (financial repression) and in turn reduces the scale of financial resources available for investment.

The second standard is how a banking system can improve the social productivities of investments in an economy. As mentioned previously, one of the big differences of banks with other financial institutions is that banks screen and monitor investment projects on behalf of their depositors. Compared with individual investors, banks can collect more and higher quality information on debtors. Furthermore, because of their credibility, participation of banks in an investment project may encourage other investors to join in. Thus, a well-functioning banking system is important in maintaining high efficiency in distributing investment funds in an economy.

Productivity of a risky investment project consists of two facts, namely, its expected return and its risk. These facts can be mathematically represented by the expected value and variation of the return. Unfortunately, these facts contradict each other in most cases. Project with high expected return usually contains high risks. According to the mean-variance theory, investors have to choose an investment portfolio with an optimal combination of expected return and risk. There is no absolute rule to determine which combination of return and risk is socially optimal, because this optimal combination is determined by the subjective utility function of return and risk of investors. However, a well-functioning banking system should be able to induce an investment portfolio that is Pareto optimal. That is, at least it is risk neutral. This means it should not prefer a project with a lower expected return compared with other projects facing similar levels of risk. Because of its publicity, governments may require a banking system to take a more conservative attitude to risk than other financial institutions.

The third standard measures the efficiency of the banking system itself. Any well-functioning industry should operate using the least social resources to satisfy the demand for their products or services from society. However, as explained previously, the banking sector should not be considered an ordinary industry, but rather more of an infrastructure. Therefore, the efficiency of the bank sector should not be measured mainly from the sector itself, but rather from its external economy to other industries. The efficiency of the banking system may spill over to other industries by lowering the pricing of transaction and financing in the production of other industries and final consumptions. However, there are cases in which the efficiency of the banking industry is mainly transformed to its own profits. In these situations, the improvement of efficiency in the banking system may not be socially beneficial. In extreme cases, when the banking sector enjoys a monopoly power, these banks may obtain high profitability through high pricing. In these cases, high efficiency of the banking sector is detrimental to the real economy.

2.2 Importance of regional banks for SMEs

The importance of regional banks in the economy can be explained by its role in financing SMEs, which have played a very important role in the economy. Berger and Udell (2002) argued that traditional lending techniques used by large banks are unsuitable for satisfying this need.

Traditional lending techniques are also called "transaction-based banking." These techniques rely heavily on hard information to form lending decisions. Hard information includes the financial statement, value of available collaterals of the borrowers, and credit scores of borrowers given by credit rating agencies inside or outside the banks. This kind of information can be quickly collected at the time of loan decision and easily transferred through the organization of lending banks.

However, these "transaction-based banking" techniques are unsuitable for SMEs. Compared with large enterprises, SMEs usually do not have formal financial accounting system. The financial statements of these enterprises are usually simple and opaque. SMEs usually do not have sufficient fixed assets as collaterals. Because of these two shortcomings, SMEs also normally cannot obtain high credit rating scores from banks or outside risk assessment agencies. All these shortcomings make obtaining credit from banks based on transactional credit techniques difficult for SMEs. Even if they can, these SMEs are much more vulnerable to "credit rationing" when banks are constrained in their ability for credit supply or wish to shrink their total credit supply because of their pessimism with the economic future. These problems also make entering the security market harder for SMEs because direct investors are even more incapable of obtaining information from SMEs.

These considerations make obtaining financing from banks difficult for SMEs. For example, in 2006, no more than 0.5 million SMEs had access to bank loans in China. Therefore, over 98% of the total 40 million SMEs in China could not obtain bank loans (Lin 2007).

The technique suitable for SMEs is relationship banking. Relationship banking is based on "soft information." This kind of information includes information on the operation history and commercial moral decency of enterprises, the characteristics of the owners of enterprises, and general business environment in which the enterprises operate. Soft information can only be obtained through long-term relations with the enterprises, their customers, suppliers, and other enterprises that have contacts with them. Soft information also requires banks to be familiar with the general regional environment. As Schumpeter (1939) stated:

"The banker must not only know what the transaction is which he is asked to finance and how it is likely to turn out, but he must also know the customer, his business and even his private habits, and get, by frequently "talking things over with him," a clear picture of the situation" (p.116).

Given its nature, soft information is usually possessed by credit officers in bank branches that have direct contact with borrowers and the regional business. Transmitting soft information through the layers of organization of the banks can be difficult. In addition, valuing such information is similarly difficult for senior officers. Thus, to offer credit to SMEs through relationship banking, it is importance to give credit officers in branches the right to authorize loans independently to customers. However, this will bring about an agency problem, because loan officers may not use their soft information for the benefit of their banks but for their own interest. Thus, banks have to build a mechanism that can supervise the behaviors of the loan officers effectively. Large banks are in a disadvantageous position under this mechanism. Given their large scale and scope of business, large banks have complex and multi-layer management structure. For the same reason, the headquarters of large banks are often located far from the regional market where an SME operates. Therefore, building an effective mechanism to monitor branch credit officers is costly for large banks. Furthermore, large banks are usually publicly listed. Thus, explaining relationship banking to their stockholders is also difficult. Given these problems, large banks mainly offer traditional transaction-based loans. Their ability to supply credit to SMEs is limited.

On the other hand, attracting funds from the regional market and contributing such fund among their nationwide branches are easy for large banks. Because of their high credibility, large banks are privileged in the competition for funds with regional banks. Thus, a phenomenon of fund outflow from less developed regions exists in many countries. Funds are collected from these regions by large banks and then distributed into large enterprises in developed regions through their web of branches.

Regional banks are much more suitable for adopting relationship banking than large banks. Similar to SMEs, regional banks themselves are usually small- and medium-sized banks. Their operation scope (called operational areas) is limited to a few regions. Given their relative small scale and geographical scope, the managerial layers of regional banks are relatively simple. The headquarters of these banks are nearer to their branches, enabling monitoring the behaviors of credit officers in their branches easier. Many regional banks are not listed in public stock markets. Therefore, these banks do not need to explain their relationship banking business to their shareholders.

Given that their businesses are focused in certain regional markets, the problem of fund outflow is not as serious in regional banks. What they obtained from the regional market will mainly be invested in the same market.

2.3 Regional financial institutions in developed countries: Shinkin banks in Japan

2.3.1 Japanese financial system and Shinkin banks

The Japanese financial institutions consist of three groups, according to their scale, operation scope, and importance in the economy. These are (1) city banks and trust banks, (2) regional banks and second regional banks, and (3) Shinyo Kinkos (Shinkin banks) and Shinyo Kumiai (credit cooperatives).

The first two are ordinary commercial banks defined as “ordinary banks” by the Department of Finance in 1968. The first group is composed of large banks. City banks are large commercial banks with nationwide operations. They are also internationally important banks. In the 1970s, 13 city banks existed. After the M&A in the 1990s, only five city banks remained from 2006 to date. Trust banks are based on the “Security Exchange Law of 1948,” which required the separation of commercial and investment banking. Apart from providing conventional commercial banking service, these banks also offer asset management services. In 1954, seven domestic trust banks existed in Japan. In 1993, the city banks and security companies were allowed to establish their subsidiary trust banks as a measure of financial liberalization. Thus, the number of trust banks increased to 16 by 2013.

The second group includes regional banks. These banks are also “ordinary banks,” but are much smaller than banks in the first group and operate over limited geographical areas. Regional banks are established under “Article 4 of the Bank Law” and members of the “Regional Banks Association of Japan.” These banks have important influence on the regional economy. The operation area of a regional bank usually covers a province (To, Dou, Fu, and Ken). From 2011 to date, Japan has 64 regional banks. The second regional banks are usually transformed from mutual banks and members of “The Second Association of Regional Banks.” These banks are smaller and less important for the regional economy. By 2013, Japan had 41 second regional banks.

The third group consists of mutual financial institutions, such as savings banks and mutual funds in the United States. Given that Shinkin banks are the object of the research in the first part of this dissertation, the third group requires detailed explanation. Mutual financial institutions originated from credit associations and first appeared in 1900 according to the “Law of Industrial Associations,” which was designed following the example of Germany. Credit associations offer financial help to farmers and SMEs, who held a weak position in the economy at the time. However, these credit associations could not accept deposits from non-members. Thus, their benefit to SMEs in the cities was quite limited. In June 15, 1951, the Japanese government published the “Law of Shinyo Kinko (Shinkin bank)” and began to establish Shinkin banks.

Shinkin banks are membership associations instead of companies. Any person who works, lives, or owns an enterprise in the area of operation of a Shinkin bank can be a member of the Shinkin bank. However, their memberships are limited to SMEs and individuals. Any individual or institution who employs more than 300 people or has a capital of more than 900 million Yen cannot be a member. Shinkin banks operate in a geographical scope similar to regional banks in the second group. These banks are also capable of attracting deposit from non-members. However, Shinkin banks are smaller than regional banks. The credits offered by these banks are mainly to its members. The loans to non-members cannot be greater than 20% of their total loans. The aim of operation of Shinkin banks is not purely profit-seeking, but to service the members and the region where they are located. The operational strategy of these banks is determined by representative committees with the rule of one vote for each member.

By the end of 2013, Shinkin banks had 9.28 million members, employed 112.525 thousand full time workers, and had a total capital of 809.7 billion Yen. These banks attracted 128060.2 billion Yen deposits and offered 64479.1 billion Yen credits.

Shinyo Kumiais operate under the "Small and medium-sized enterprise cooperative saving insurance corporation law of 1949." These cooperatives are smaller than Shinkin banks. The operations of these cooperatives are limited to the counties where they are located and are not allowed to attract deposits from non-members.

From the end of 2010 to the end of 2013, the number of Shinyo Kumiais declined from 323 to 159.

2.3.2 Evolution of the Japanese financial system and Shinkin banks

Until the mid-1980s, a unique characteristic of the Japanese financial system is the dominant role of bank loans as sources of external financing for enterprises. By the mid-1980s, bank loan accounted for over 80% of the total corporate finance. The remaining 20% were mainly from new stock issuances (Imai and Takarabe 2011). Bond issuance only played a minimum role in corporate financing.

Another characteristic of the Japanese financial system before the mid-1980s is the strong relationship between government, banks, and corporations. For a long time, the Japanese banking system was strictly controlled by the government. The operations of different kinds of financial institutions were clearly divided. Loan decisions of banks were guided by the government to suit its industry policies. Interest rates were deliberately kept at a level lower than the market level. In return, the government tacitly offered guarantees of bail outs if the banks were in trouble. The relationship between banks and corporations was also very strong. The most important aspect was the main banking system existed between banks and corporations. This system enabled banks to build a long term relationship with corporations through financial transactions (deposit, loan, and foreign exchange, etc.) and inter-shareholding. A close relationship also existed among banks themselves through inter-shareholdings. This unique financial system helped Japan to achieve a high economic growth rate after World War II. However, when the Japanese economy became mature and entered the relatively slow but stable economic growth stage, the shortcomings of this non-market oriented financial system began to show and caused inefficiency and a huge volume of NPLs in the banking system.

From the mid-1980s, influenced by the international trend of financial liberalization，the Japanese government began to liberalize the financial system. The first plan was to remove the deposit interest rate ceilings. From 1985 to 1996, the deposit interest rates were gradually liberalized. From 1985, the Japanese government relaxed its control over corporate bond issuances and reformed the stock exchange market. Foreign exchange control was also removed. These changes enabled seeking direct financing from domestic and foreign security market easier for large corporations. Since then, large corporation gradually increased the proportion of bond issuances in their external financing. Thus, the dependence on loans of large Japanese corporations decreased. From the mid-1980s to 1995, the ratio of bank loans to total external financing of large corporations declined from 80% to 75% (Imai and Takarabi). However, the reform was generally non-radical, unbalanced, and it also brought out negative side-effects. Given that the demand from large corporations declined, banks had to search for replacement customers for loans. Many increased their loans in the booming mortgage market. This stimulated the formation of bubbles in the real estate market.

In 1996, the Japanese government launched a more radical financial liberalization program. The newly elected Prime Minister, Ryutaro Hashimoto, announced a plan called "Reform of Japan's Financial System - For the Rebirth of the Tokyo Market in 2001." The program was called the Japanese version of "Financial Big Bang" because of its similarity with the British plan in the 1980s. It aimed to raise the Japanese financial system to international standards by 2001. According to the program, the separation of operation between different kinds of financial institutions would be gradually cancelled. Banks, insurance, and security companies would be permitted to directly enter into the business areas of each other. The program also planned to raise the accounting system of financial institutions to the international standard (Anderson and Terry L. Campbell 2000). The government forced banks to increase transparency of their accountings, especially the accounting of NPLs. Loans were now classified into four categories according to the possibility of repayment. The program also instructed banks to save adequate loan loss reserves and establish plans to reduce bad loans. A consolidated accounting system replaced the system where only the financial status of parent corporations was revealed.

However, the next year after the declaration of the "Big Bang" program (1997), the real estate market bubble burst and triggered a severe financial crisis in Japan. In 1997, three famous large financial institutions (Sanyo Securities, Hokkaido Takushoku Bank, and Yamaichi Securities) declared bankruptcy. In the next year, two other large financial institutions (Long-Term Credit Bank of Japan and Nippon Credit Bank) were placed under the protection of the government. The NPLs of the remaining banks increased sharply. Japanese large banks had to pay 100 basis point premiums when borrowing abroad. The severe conditions forced banks to cut their loan supply; hence, a "credit crunch" prevailed in the Japanese economy. Japan entered into a period of long term recession.

The Japanese government had to focus their attention on addressing these problems in the banking system. In February 1998, the government established a "Financial Function Stabilization Plan." The plan offered 30 trillion Yen to Deposit Insurance Corporations (DIC). The fund was used to provide full deposit protection and supplement capital of large city and trust banks (Montgomery and Shimizutani

2009). A few institutional changes were also made. In June 1998, the responsibility of regulating financial institutions was separated and transferred from the Ministry of Finance to the then newly established "Financial Supervisory Agency (FSA)." This agency was directly responsible to the "Financial Reconstruction Committee," which was also a newly established agency in December 1998. However, this reaction of the Japanese government was widely criticized as being too late and too slow. The volume of financial support was also deemed too small.

After entering the 21st century, the Japanese government began to deal with the NPL problem in the banking system more seriously. In October 2002, the government published the "Financial Revival Program" (Takenaka program). Compared with the policies in the latter part of the 1990s, the program declared a stricter schedule to solve NPL problems. According to the program, the NPL ratio should be reduced by half by the end of 2004. The schedule was conducted as planned. By the end of FY 2004, the problem of NPL in Japan was basically settled. Thus, by December 2004, the "Financial Revival Program" was substituted by the "Financial Reform Program." The focus of the financial policy was transferred from the stability of the financial system to its vitality. Through this program, the Japanese government significantly restructured its financial system. In November 2005, the government announced an amendment to the bank law. This amendment greatly liberalized the operation of financial institutions. The requirement of 100% ownership for bank subsidiaries was cancelled. Strict separation of business areas between different kinds of financial institutions was withdrawn. The regulation for the banking agency was greatly relaxed.

Meanwhile, because of the key role of regional financial institutions in the economy, from 2000, the government also began to design programs different from those for major city banks to improve the operations of regional financial institutions. In August 2004, the government published a "Financial Function Strengthening Act." The act allowed the government to fill in public funds to financial institutions without considering its importance in the financial system. Thus, using public funds to assist regional financial institutions became possible. In March 2005, the "Action Program

Related to the Improvement of the Function Strengthening of Regional Financial Institutions" was published. The act offered financial assistance to regional financial institutions. From March 2003 to September 2009, 10 regional financial institutions received financial funding from the government with a total value of 282.5 billion Yen.

Moreover, to strengthen the financial system, the government reinforced the support for the M&A between financial institutions in Japan, both among similar and different kinds of institutions. In December 2002, the government published the "Act of Special Measure to Promote the Merging and Restructuring of Financial Institutions." This act triggered a wave of M&A among Shinkin banks and between Shinkin and regional banks or other types of credit associations (such as Shinyo Kumiais). After the M&A, the number of Shinkin banks decline drastically from 396 by the end of 2000 to 267 by the end of 2013.

2.4 Regional financial institutions in developing countries: City commercial banks in China

2.4.1 Outline of the Chinese banking system and Regional financial institutions in China

Chinese banks can be divided into three groups. The first group consists of five large banks that have nationwide operations. Four of them are state owned commercial banks[2] (the Big Four) and the other one (Bank of Communication or BankComm) is a state-controlled commercial bank. By the end of 2012, they still accounted for approximately 47.34% of the total bank assets in China.

The second group includes 12 joint-equity banks, which were established jointly by local government, state owned enterprises, and private investors. These banks have been established since 1987. In 2012, these banks became important credit suppliers

[2] Agriculture Bank of China (ABC), Bank of China (BOC), China Construction Bank (CCB), and Industrial and Commercial Bank of China (ICBC).

in China and accounted for approximately 16.22% of the total bank assets in China.

The third group includes urban commercial banks and rural commercial banks. These banks are regional banks operating in certain urban and rural areas, respectively. Even before the economic reform, apart from the People's Bank of China (PBC), which was the only bank at the time, rural credit cooperatives existed. These cooperatives were collectively owned by financial institutions operating in rural areas. In 1979, the Chinese government also established similar urban credit cooperatives in urban areas. In 1994, urban credit cooperatives were restructured to become urban cooperative banks. In 1998, these banks were changed to urban commercial banks through transformation, merging, and acquisition. In 2001, rural credit cooperatives were similarly changed into rural cooperative banks and rural commercial banks. By the end of May 2012, China had 137 urban commercial banks, 155 rural commercial banks, and 210 rural cooperative banks. By the end of 2012, the total asset of urban commercial banks reached 9984.5 billion RMB Yuan, making up for approximately 8.81% of the total bank assets in that year.

Most of the urban commercial banks and rural commercial banks are small banks operating in a single municipality (Shi). However, in recent years, a few regional banks have rapidly increased their scales. The operating areas of these banks exceed the boundary of the municipality they originally located. A few of these banks (such as Bank of Beijing and Bank of Hangzhou, etc.) have become as large as joint-equity banks. However, the operation areas of these banks are still constrained in several large and middle cities. By 2011, China had 27 urban commercial banks with total assets exceeding 100 billion RMB Yuan. Among these urban commercial banks, three banks have already been listed in the stock market. Another 11 banks are waiting to be listed.

2.4.2 Evolution of the Chinese banking system and regional financial institutions

Initially outlining the banking system in China from 1949 to 1979 and its characteristics will be helpful in understanding the background of bank restructuring in China.

After the completion of the "Program of Socialist Transform" in 1954, a socialist (central planning) economic system was established in China. Thus, a new banking system was established to suit the new economic system.

One of the essential characteristics of central planning economies is that almost all economic activities are tightly controlled by the central government. Neither commodity nor financial market existed. The Banking system was merely a tool of central planning. Before 1979, only one bank actually existed, namely, the PBC. The PBC played both the role of a central bank and a commercial bank.

Fixed investment decisions were not made by managers of the firms, but by the central government according to the fixed investment plan. Bank loan decisions were also not made by the bank managers themselves, but by the central government according to a credit plan, which was a subordinate to the fixed investment plan. However, both firms and banks would not be responsible for the consequences of the investments. Under these circumstances, the managers of firms and the local authorities unsurprisingly had a strong tendency to try to obtain investment from the central government as much as possible, because they would lose nothing if the investment went wrong. However, these managers would greatly benefit if the investments were successful. For the same reason, bank managers also preferred to expand their loans as large as possible. This phenomenon is the so-called "soft budget restrictions." Thus, during that time in China, a strong tendency for economic expansion was observed. Overheating and inflation were always among the key problems. However, since banks only provided working capitals and these loans were based on "real bill" doctrine, NPLs were unlikely to become a serious problem.

In 1979, China began the economic reform. The government realized that the old banking system was no longer suited for the new market-oriented economic system and had to be reformed. From 1979 to 1983, the reform focused on the separation of central bank operations and commercial banking. Since 1981, four specialized banks have been established or re-established. The PBC shifted its commercial bank business to the "Big Four" and became a sole central bank.

After the separation of the central bank and commercial banks, the Chinese government began to gradually substitute budgetary allotment by bank loans. This policy resulted in the rapid decrease of the proportion of fixed investment financed by the government. Almost all fixed investments were financed by bank loans except for a few national "key" projects. The so-called "substitute budgetary allotment by bank loans" policy made firm managers more conscious of the cost of investments because budget allotment was free but bank loans had to be repaid with interest. However, loan decisions were still not independently decided by bank managers based on credibility but directed by the government. The loans directed by the central government were called "Policy loans." Up to 1993, it still accounted for approximately 15% to 45% of the total loans provided by the "Big Four" (Xiaoping Xu 1997). To allow the "Big Four" to get rid of "policy loans" and make them true commercial banks, three policy banks were established in 1993. These banks took over the policy loans from the "Big Four." However, in practice, many loans offered by special banks were still directed by the government. In many cases, bank loans were still simply replacements for the budget allotment. Thus, the "soft budget" problem still existed after the reform. However, most of these policy loans were mid-term credit for fixed investments. Hence, the risk of default was very high. This phenomenon was the main reason why NPLs accumulated rapidly in China in the 1990s. For regional banks such as city and rural commercial banks, the problem was even more severe. Given that these banks were more strongly influenced by the local government, they were more eager to offer credits to support the regional economic growth. These banks offered high interest rates to attract deposits and offered loans to highly risky investment assets. A few of these banks encountered huge troubles in the 1990s and had rescued by the government or even had to declare bankruptcy.

In 1998, the central government of China began to notice the severe situation of banks. The central government launched a more fundamental bank-restructuring program. First, the PBC cancelled the quantity control of bank loans and substituted it with asset-debit management policy. In 1996, China became a formal member of Basel Accord. In 2003, the Chinese Banking Regulation Committee (CBRC) was established as bank supervisor. In 2004, the CBRC published the "Measure of Capital Adequacy Ratio Management for Commercial Banks." According to this directive, from then to the end of 2006, the CBRC gradually implemented the CAR into the banking system. This directive was a mixture of Basel I and Basel II, which were newly published at that time.

In accordance with the publication of Basel III, in June 2012, CBRC once again published a directive, "Measure of Capital Management for Commercial Bank." According to this directive, for nonsystematic domestic important banks (most of the regional banks are included in this group), core first tier capital CAR (including common stock), first tier capital (including common stock, general risk provision, and undistributed surplus), and total capital CAR (including first tier capital plus supplementary capital, such as long term debt) would reach 5.5%, 6.5%, and 8.5%, respectively, by the end of 2013. By the end of 2018, these requirements would be increased to 7.5%, 8.5%, and 10.5%, respectively. For systematically important banks, a further 1% core capital is required. Thus, by the end of 2013, their minimum CAR requirements were raised to 6.5%, 7.5%, and 11.5%, respectively, and they would be increased to 8.5%, 9.5%, and 11.55%, respectively by the end of 2018.

There are signs that this policy is effective in curbing the impulse of banks to oversupply credits. Many banks are rumored to be in shortage of capital. They are lined to apply for listing in domestic markets and oversea stock markets. Those who are already listed in stock markets are frequent issuers of new stocks. In doing so, they are criticized for damaging the benefits of the old owners of their stocks. Capital supervision effects on loan supplies of Chinese regional banks are discussed in detail in Chapter 7.

Second, the Chinese government began to deal with the severe problems of NPLs in the banking system. PBC changed the classification method of bank loan quality. Before 1998, China used a four-category division method to classify bank loan quality. The shortcoming of this division approach was that bank loans were based on payment status and not on recovery possibility. This made possible for banks to conceal bad loans by providing new loans to borrowers who actually had lost the ability to repay the loans. Besides, many loans in China were so-called bullet loans," which need not repay the principal until the end of the term. Therefore, if a borrower went bankrupt, the loans offered to him would not immediately become an NPL until the end of its term.

In 1998, with the help of United States, Chinese government adopted a new credit classification system based on repayment possibility. This new system classifies bank loans into five categories: normal, needs special attention, abnormal, doubt, and loss.

The government also offered fiscal help for the banks while putting more pressure on the banks to control NPL growth. In 1998, financed by a special issue of treasury bill, the Chinese government injected 279 billion RMB Yuan into the "Big Four." The money caused the core CAR ratio of the "Big Four" to increase by 4.7% on average, which is above the minimum requirement of Basel Accord at least in book value. In 2003, the Chinese government once more provided 45 billion US dollars to the two of the "Big Four" (Bank of China and Construction Bank of China) through foreign exchange reserve so that they can be listed in the stock market.

In 1999, the Chinese government established four asset management companies (AMCs) to help the "Big Four" in dealing with NPLs. One bank was assigned for each AMC. By mid-2000, the transfer of NPLs was completed. A total NPL worth 1400 billion RMB Yuan were transferred to the four AMCs.

After entering the 21^{st} century, the Chinese government acknowledged the importance of SMEs in the economy and the role that regional banks played in SME development. In 2003, the government issued a "small- and medium-sized enterprise promotion law." The law was announced to offer several types of financial support, including credit supply and credit guarantee services to SMEs. However, the law did provide detailed policy explanations and complementary laws. Thus, it lacked practical maneuverability and forcibility.

From 2008, PBC, China Banking Regulatory Commission (CBRC), China

Securities Regulatory Commission, and China Insurance Regulatory Commission announced several policies to offer practical financial supports for SMEs. In 2008, CBRC asked commercial banks to set separate plans for composite credit lines for SMEs. These credit lines should be examined, accounted, and valued separately. For the interest rate and other financing costs, the government allowed banks to charge a reasonable higher interest rate to credits offered for SMEs so that the banks could cover the extra risks of these credits. The government gradually increased the range of loan interest controlled by the government to enable the application of these policies. In April 2014, PBC announced a targeted required reserve rate (RRR) cut. The RRR of rural commercial banks were lowered by 2 percent point, and the RRR of the rural cooperative banks were lowered by 0.5 percent point. In June 2014, PBC further increased the objects of target RRR. All banks whose new credits to agriculture, small enterprise, and microenterprise accounted for over 50% of their total new credits supplied last year (many regional banks met the requirement) could enjoy a 0.5 percent point cut of RRR. By contrast, in 2011, the government forbade banks from taking commitment and management fees for loans to SMEs to reduce financial cost of SMEs. Financial consultant fees were also restricted.

Chapter 3 Theories on measuring efficiency and productivity changes

In this dissertation, I used a nonparametric approach called Data Envelopment Analysis (DEA) to estimate the efficiency of Shinkin banks in Japan. Based on the estimates, a measure called Malmquist index is also calculated to estimate the productivity changes of these banks. A major difference of the analysis in this dissertation from the analysis in the existing literature is that a bootstrapping approach is used to compute the robustness and statistical significance of the estimates. The results are shown in detail in the next chapter. As a theoretical basis, the theoretical background and technique details of related concepts, such as DEA, bootstrapping, and Malmquist index, are explained in this chapter.

3.1 Measurement and estimation of efficiency

Efficiency is the relative performance of a decision-making unit (DMU) compared with its potential performance. Koopmans (1951) was the first to define the concept of production efficiency, that is, the production of a DMU is efficient if and only if, under the current input constraints, no way to increase one unit of a product exists without the need to reduce the production of at least one of the others. In duality, it can also be defined as impossible to reduce one input without the need to increase other inputs for a given level of outputs. This concept is similar to that of Pareto efficiency, which is taught in standard economics books.

3.1.1 Measurement of efficiency

The simplest measurement of production efficiency is some ratio in the form of output per unit of input, such as the following:

$$\text{Efficiency index} = \frac{\text{output}}{\text{input}}$$

This simple ratio is a fine measurement of efficiency in the case of one output–one input situation. When there are multiple inputs and outputs, production efficiency can still be partially reflected; however, it no longer reflects the production efficiency as a whole (called the total factor productivity, TFP) (Cooper, et al. 2006). If the prices of inputs and outputs are available, we still can use simple indices, such as profit rate or average cost, as a measure of efficiency. However, prices of outputs and inputs are not always available, particularly in nonprofit organizations. Even in some enterprises, several prices (such as the price of labor inputs or capital in some industries) are also not easy to calculate. In these situations, a single index of efficiency cannot be obtained. Even if a single index of efficiency is obtainable, we cannot detect the source of inefficiency. For example, if an index of efficiency indicates that a DMU is inefficient, we cannot know whether it is because of waste of inputs, non-optimal mix of product or distribution of inputs, or lack of economic scale.

Debreu (1951)and Farrell (1957) are the first to introduce the theoretical background of efficiency measurement in multiple inputs and outputs situations. The Debru and Farrell measurement can be defined from the direction of either input or output. From the direction of input, the efficiency of a DMU is defined as the maximum equal-proportional (radial) reduction of all inputs possible while maintaining the same level of production. In the output direction, efficiency is defined as possible maximum radial augmentation of all outputs under the constraint of inputs.

Mathematically, we can define the production set Ψ^t as:

$$\Psi = \left\{(x, y)\Re^{N+M} \mid x \text{ can produce } y\right\}$$

where $x \in \Re^N$ is the N dimension vector of inputs, and $y \in \Re^M$ is M dimension vector of outputs. Farrell input-oriented radial efficiency measurement θ is defined as:

$$\theta^*(x, y| \Psi) = \inf\{\theta>0|| (\theta x, y) \in \Psi\}$$

θ is clearly low bounded by 0 and up bounded by 1.The higher the θ, the higher efficient the point of (x_0, y_0) becomes. Correspondingly, those DMUs with $\theta = 1$ establish an efficient set (frontier) of inputs $\partial X(y)$.

Similarly, we can define a Farrell output-oriented radial efficiency measurement β as follows:

$$\beta^*(x, y| \Psi) = \sup\{\beta>1| (x, \beta y) \in \Psi\}$$

An efficient set (frontier) of outputs $\partial Y^t(x)$ consists of DMUs with $\beta = 1$ is also obtained.

Both input and output-oriented efficiency can be measured under the assumption of constant return to scale (CRS) or variable return to scale (VRS). The frontier and input- and output-oriented Farrell efficiency measure are explained by Figure 3.1.

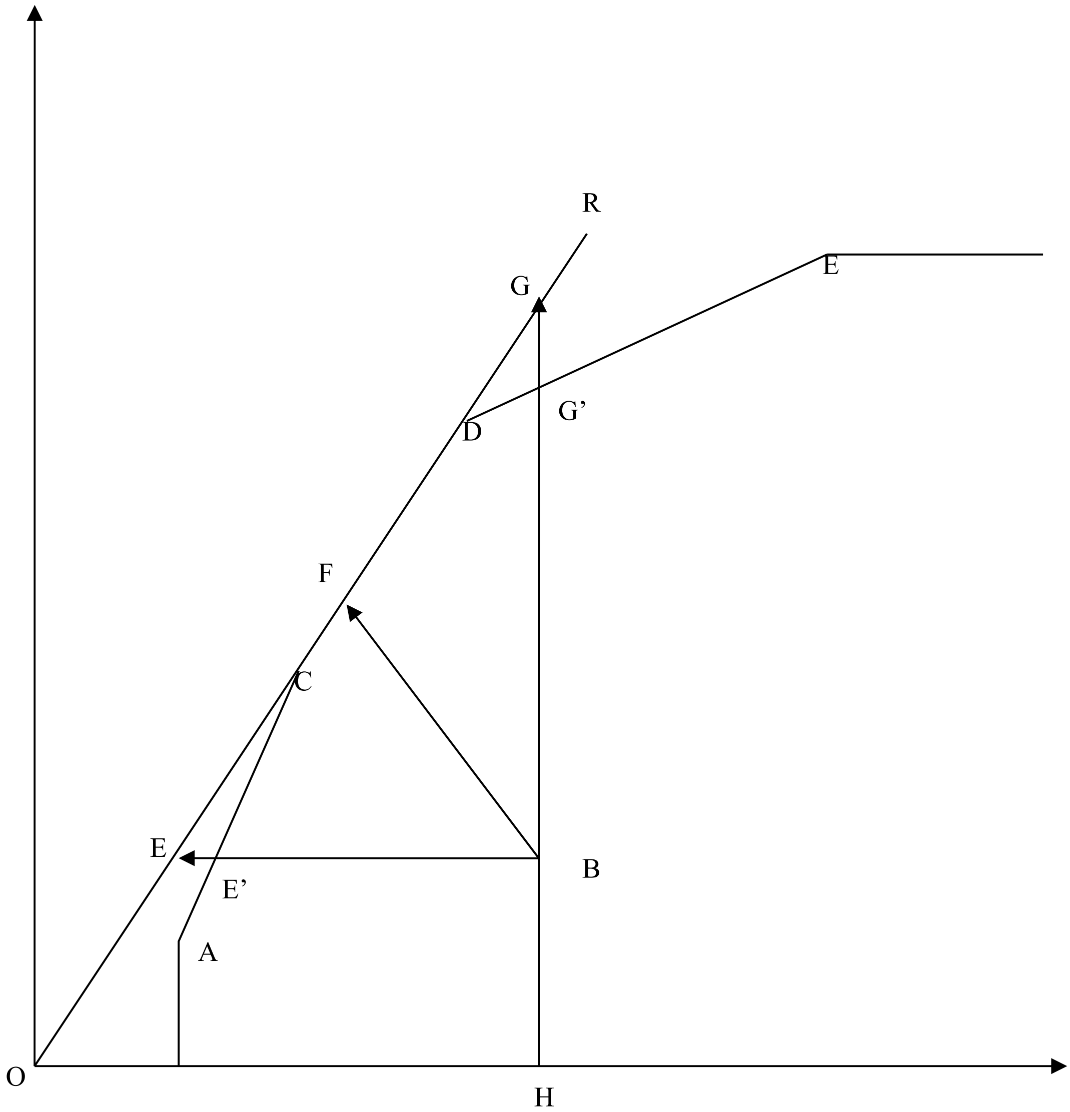


Figure 3.1: Farrell efficiency measurement along different directions

In Figure 3.1, the production sample is constituted by five DMUs: A, B, C, D, and E. Among them C and D are efficient under the assumption of CRS. A, C, D, and E are efficient under the assumption of VRS. CRS frontier is the ray OR that connects C and D. VRS frontier is the line sections connecting the efficient points of A,C,D, and E. Thus, under the assumption of CRS, the input-oriented distance of inefficient point B is $\theta_{CRS} = BE/HO$. Its output-oriented distance is $\beta_{CRS} = HG/BG$. By contrast, under the assumption of VRS, the input-oriented distance of inefficient point B is $\theta_{VRS} = BE'/HO$, and its output-oriented distance is $\beta_{VRS} = HG'/BG$.

It is obvious $\theta_{CRS} = BE / HO = BG / HG = 1 / \beta_{CRS}$. Thus, under CRS assumption, both the input-oriented and output-oriented efficiency measurements are the same for any DMUs. Therefore, it makes no difference whether you choose the input-oriented measurement or output-oriented measurement. However, under the assumption of VRS, the results may be significantly different, particularly for DMUs located at the extremes of the production set. The input-oriented distances of DMUs with small inputs lying near the low-left side (the steep part) of the frontier, which corresponds to the part of the frontier with increasing return, are larger than output-oriented distances. By contrast, the input-oriented distances of DMUs with large outputs lying near low-left side (flat part) of the frontier (corresponding to the part with decreasing returns) are smaller than output-oriented distances. The difference depends on the curvature of the VRS frontier and the distribution of the production set.

The choice of direction may also influence the return to scale. Figure 3.2 illustrates this situation, which is similar to that by Fukuyama (1996). In Figure 3.2, as in Figure 3.1, CRS frontier is represented by line OR, and VRS frontier is represented by curve ACDE. In VRS frontier, line AC represents increasing returns to scale (IRS) frontier, line CD represents CRS frontier, and line DE represents decreasing return to scale (DRS) frontier. For DMUs located in area II, IRS will be shown regardless of which direction is chosen. Similarly, DMUs located in area CC always show CRS, whereas DMUs located in area DD always show DRS. However, in area IC, DMUs show IRS if efficiency is measured in the input orientation, but CRS if measured in the output orientation. In area ID, DMUs are IRS in input orientation, but DRS in output orientation. In area CD, DMUs are CRS in input orientation, but DRS in output orientation.

At present, no theoretical foundation that can guide the choice between input or output directions exists. Thus, measuring efficiency under other directions would be helpful in obtaining an unbiased measurement of efficiency.

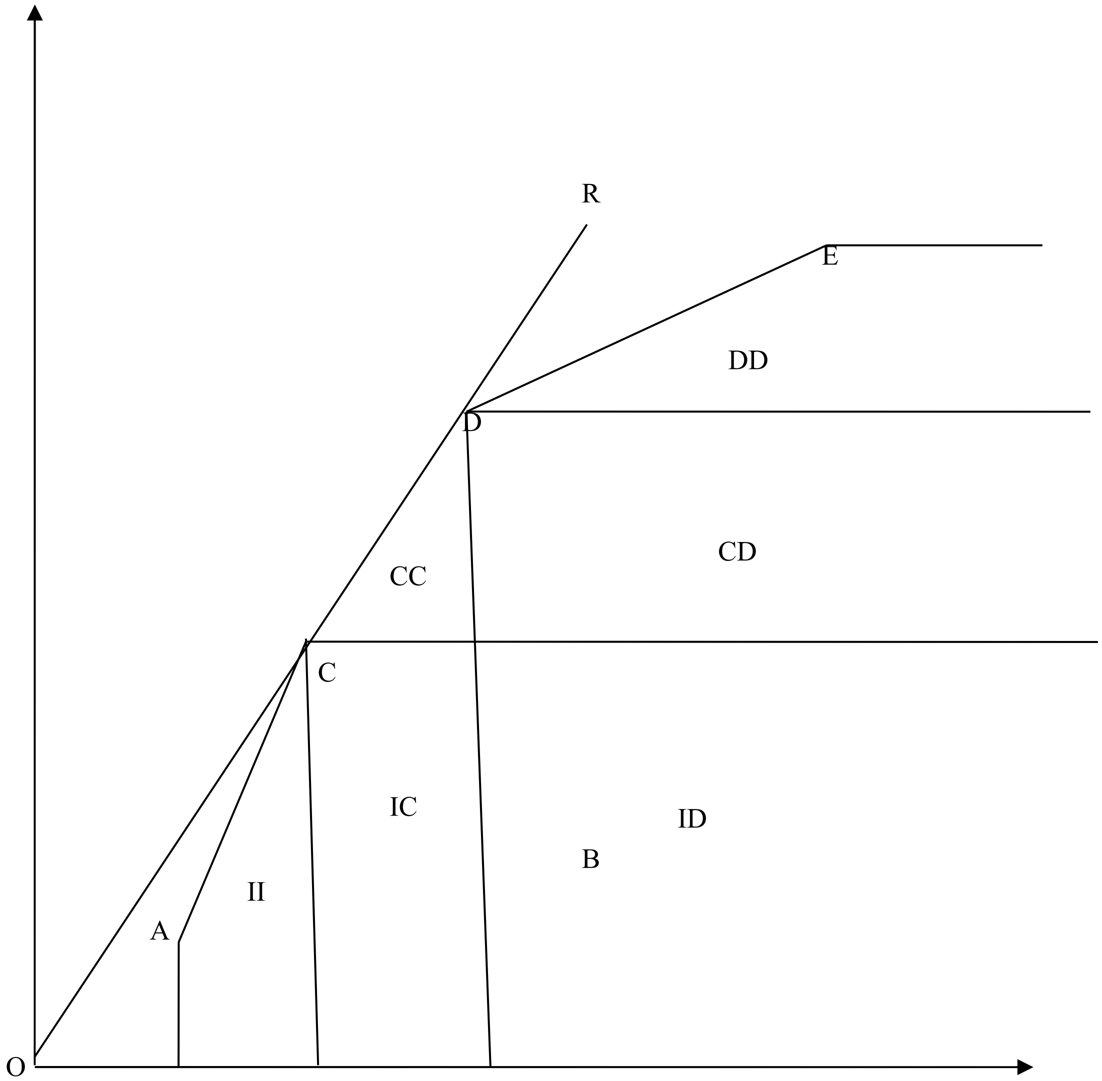


Figure 3.2: Influence of orientation on the measurement of return to scale

One of the most widely used measurements of this kind similar to the Farrell radial efficiency measurement is the so-called hyperbolic-oriented measurement. The hyperbolic-oriented efficiency measurement is defined as the proportion needed to push a production point to the frontier by simultaneously reducing inputs and increasing outputs under the constraints. That is,

$$\phi^*(x, y | \Psi) = \inf\{\phi > 0 | (\phi x, y/\phi) \in \Psi\}$$

In Figure 3.1, the hyperbolical-oriented Farrell efficiency measurement for point B is represented by line BF, which is the line originating from B to the frontier along the 45° angle. Given that F is located in line CD, which is the CRS part of the frontiers, the CRS and VRS hyperbolical-oriented Farrell efficiency measures are the same.

Another widely used efficiency measurement is Shephard distance, which was named after its founder R.W.Shephard (1953). Shephard distance is the relative distance from the point of a DMU in the output–input space to the "frontiers" of production. Shephard distances can also be measured along the direction of inputs or outputs. The production set Ψ is defined as the same in the case of Farrell distance. Thus, Shephard input-oriented distance D_{input} is as follows:

$$D_{input} \equiv \sup\{\delta > 0 \mid (x/\delta, y) \in \Psi\}$$

where $1/\delta$ is the proportion needed to scale down x_0 to push point (x_0, y_0) along the x axis to the frontier .

Similarly, the Shephard output-oriented distance D_{output} is defined as follows:

$$D_{output} \equiv \inf\{\varphi > 0 \mid (x, y/\varphi) \in \Psi\}$$

where $1/\varphi$ is the proportion needed to scale up y_0 to push the point (x_0, y_0) along the y axis to the frontier.

We can also define a Shephard hyperbolic distance D_{hyper} as:

$$D_{hyper} = \sup\{\gamma > 0 \mid (x/\gamma, \gamma y) \in \Psi\}$$

where D_{hyper} is the distance of point (x_0, y_0) in time t to the efficient input frontier in time t ($\partial X(y)$) and γ is the factor needed to push the point (x_0, y_0) to the frontier by scaling down x_0 and scaling up y_0 simultaneously.

It is easy to prove that the Shephard oriented distances are just the reciprocal of the Farrell radial efficiency measure. That is: $\delta = 1/\theta$, $\varphi = 1/\beta$, and $\gamma = 1/\phi$.

Similar to the Farrell efficiency score, Shephard distance is also estimated either under the assumption of VRS (D_{VRS}) or CRS (D_{CRS}).

3.1.2 Scale economy and decomposition of efficiency

From Figure 3.1, we see that DMUs located on line sections AC and DE are efficient according to the VRS frontier. However, they are inefficient according to the CRS frontier. This situation occurs because although these DMUs operate efficiently under fixed production scales (i.e., pure efficiency), they do not operate under optimal scale (i.e., economy of scale). We can improve the total efficiency of these DMUs by shifting production combinations from line section AC or DE to CD along the VRS frontier.

The difference is due to economy of scale. We can measure the economy of scale for a bank as the ratio of these two measures of efficiency:

$$SC = \frac{D_{CRS}}{D_{VRS}} \qquad (3.1)$$

where SC is the measurement of scale, and D_{VRS} and D_{CRS} are the distance estimated under the assumption of VRS and CRS, respectively. For Shephard input- and hyperbolic-oriented distance, $scale \geq 1$.

Thus, we can decompose the efficiency measurement into two components: pure efficiency and economy of scale. We define pure efficiency as the efficiency measured under the assumption of VRS. The relationship among these three concepts can be expressed as follows:

$$D_{CRS} = D_{VRS} \times SC \qquad (3.2)$$

However, D_{CRS} only shows whether a DMU is at CRS, but does not indicate if DMU is at IRS or DRS. For no matter a DMU is at IRS or DRS, its scale score is larger than 1 for the input or hyperbolic-oriented distance.

We need to construct a frontier under the assumption of non-increasing return of scale (NIRS) to determine whether a DMU is at IRS or DRS. In Figure 3.1, this frontier is constructed by the line section OCDE. Then, we can obtain another estimation of score:

$$SC^{*} = \frac{D_{NIRS}}{D_{VRS}} \qquad (3.3)$$

D_{NIRS} is the distance measured under NIRS assumption. By construction, for input- or hyperbolic-oriented distance, if $scale > 1$, but $scale^* = 1$, a DMU is at DRS; if $scale > 1$, and $scale^* > 1$, a DMU is at IRS.

3.1.3 Estimation of efficiency

Roughly speaking, two approaches have been developed to estimate the preceding radial production efficiency measurement. The first is the parametric approach, which was first introduced by Aigner, Lovell, et al. (1977) and Meeusen and Broeck (1977). The method assigns some function form to production technology and derives a parametric econometric model from it. Efficiency is estimated by decomposing the error term of the model into a white noise term and an item that reflects the inefficiencies of individual firms. In addition to the function form, the prices of the inputs and outputs also have to be assumed in parametric models.

Another efficiency measurement approach is called nonparametric approach. Data envelopment analysis (DEA) is one of the most widely used among them. DEA identifies efficient DMUs through linear or nonlinear programming. The production frontier is the convex combination of optimal points. Efficiency measurement obtained through DEA is similar to that obtained through Farrell efficiency measurement.

Suppose we observe a production sample $Q_n = (x_i, y_i),\ i = 1 \cdots n$ from the production set Ψ defined as in Section 3.1. If we assume constant scale of return (CRS), we can calculate the input-oriented efficiency score for a fixed point (x_0, y_0) by solving the following linear program (CCR model):

$$
\begin{aligned}
(CCR_0) \qquad & \min_{\theta,\lambda} \quad \theta_{CRS} \\
\text{Subject to} \quad & \theta_{CRS} x_0 \geq X\lambda \\
& y_0 \leq Y\lambda \\
& \lambda \geq 0
\end{aligned}
$$

A radial efficiency measure for a fixed point $\theta_{CRS}(x_0, y_0)$ is obtained by solving the preceding linear programming problem. By calculating θ_{CRS} for every point in the sample, we can obtain an input efficiency frontier $\partial X(y)$, which is the linear combination of the efficient points (x, y) (with $\theta = 1$) and the attainable production set Ψ estimated by DEA:

$$\Psi_{CRS}^{DEA} = \left\{(x, y)\Re^{N+M} \mid y \leq Y\lambda, x \geq X\lambda,\ \lambda \geq 0\right\}$$

We can obtain output-oriented efficiency scores as well as their corresponding efficient output frontier $\partial Y(x)$ and attainable production set $\Psi(x, y)$ by using similar linear programming models as in input-oriented cases. The following is the corresponding linear programming:

$$\begin{aligned} &(CCR_0) && \max_{\beta,\mu}\ \beta_{CRS} \\ &\text{Subject to} && x_0 \geq X\mu \\ & && \beta_{CRS}\, y_0 \leq Y\mu \\ & && \mu \geq 0 \end{aligned}$$

We can also calculate the hyperbolic-oriented efficiency for a fixed point (x_0, y_0) by solving the following programming:

$$\begin{aligned} &(Hyperbolic_0) && \max_{\gamma,\eta}\ \gamma_{CRS} \\ &\text{Subject to} && x_0 / \gamma_{CRS} \geq X\eta \\ & && \gamma_{CRS}\, y_0 \leq Y\eta \\ & && \eta \geq 0 \end{aligned} \tag{3.4}$$

DEA efficiency scores can also be calculated under the VRS assumption. The only difference between the CRS and VRS models is that the latter includes a new constraint $e\lambda = 1$ (e is a N dimension vector of 1). For example, in hyperbolic-oriented models, the following is the program (BCC model) used to obtain the score for a fixed point (x_0, y_0) under assumption of VRS:

$$
\begin{aligned}
&(Hyperbolic_0) \qquad \max_{\gamma,\eta} \quad \gamma_{VRS} \\
&\text{Subject to} \quad x_0 / \gamma_{VRS} \geq X\eta \\
&\qquad\qquad\quad \gamma_{VRS} y_0 \leq Y\eta \\
&\qquad\qquad\quad e\eta = 1 \\
&\qquad\qquad\quad \eta \geq 0
\end{aligned} \tag{3.5}
$$

Unlike traditional input- or output-oriented models, programming in the case of hyperbolic distance is not linear. In the case of CRS, the input- and output-oriented distances are the same, so it is easy to prove that $\gamma_{CRS} = \theta_{CRS}^{1/2} = \beta_{CRS}^{1/2}$ (Fare, 1985); therefore, we can easily obtain γ_{CRS} by estimating the input- or output-oriented distance. In contrast, the solving the case of VRS is considerably complex. However, with the help of a computer, the programming may also be solved with numerical algorithm (Wheelock and Wilson, 2009).

Compared with the parametric approach, one advantage of DEA is that it does not assume any functional form for the production function, thereby avoiding the difficulty of specifying a correct functional form. Wilson (2008) proved that when the size of the DMUs is widely distributed, the translog function form (a most widely used production function) is specified incorrectly. Another strong point of DEA is that it does not require any information on prices of outputs and inputs, which are sometimes difficult to gather or at times do not even exist.

However, the traditional DEA also has several shortcomings. Unlike the parametric approaches, the traditional DEA does not account for random effects in models, because it assumes that no random factors will temporarily affect efficiency scores. Similarly, measurement errors are also ignored. Nonparametric models, such as DEA, are known to be more sensitive to outliers than parametric models. More importantly, since traditional DEA models are deterministic in nature, it is impossible to test the significance of DEA estimates statistically.

3.1.4 Bootstrapping the efficiency score

As mentioned earlier, traditional DEA efficiency scores are deterministic in nature. No random effects or errors are assumed. Several researchers have recently tried to overcome this weakness. For example, Kneip et al. (1998) investigated the asymptotic characteristics of the input radial score estimator $\hat{\theta}$ for multiple output and input cases. They found that $\hat{\theta}$ is a consistent estimator for θ, and the convergence rate is $n^{\frac{2}{p+m+1}}$, that is:

$$\hat{\theta}(x,y)-\theta(x,y)=O_P(n^{-\frac{2}{p+m+1}})$$

where p and m are the number of output and input, respectively. This equation indicates that with the increase of input and output numbers, the convergence rate decreased exponentially, indicating that a large number of samples are necessary to obtain a reasonable estimate. Even in a single output and input case (p+m =1), it will converge to the actual θ at a rate smaller than $\sqrt{n}$ in the traditional econometric model. This situation is called the "curse of dimensions," which often encountered in nonparametric estimates.

When the exact distribution form of an estimator is unknown, bootstrapping becomes an appealing instrument to analyze the statistical properties of the estimator. Simar (1992) is the first to introduce bootstrapping method into the production frontiers estimation.

In the case of hyperbolic-oriented efficiency estimator $\hat{\gamma}$, bootstrapping means generating B new samples of production sets $Q_{ij}^{*}=(x_{ij}^{*},y_{ij}^{*}),\ i=1\cdots n, j=B$ by repeatedly generating new samples from the original sample B times (data generating process), and then calculating the corresponding hyperbolical distance for each DMU ($\hat{\gamma}_{ij}^{*},\ i=1\cdots n, j=1,\cdots,B$). We can obtain several statistical properties of $\hat{\gamma}$ through the distribution of $\hat{\gamma}^{*}$. These properties enable us to draw inferences on $\hat{\gamma}$. In our case, we can calculate confidence intervals or test several hypotheses on $\hat{\gamma}$.

The simplest way of bootstrapping is to repeatedly draw items with replacements uniformly from the original sample. The advantage of this simple (naïve) bootstrapping is that it does not require estimation of the probability density function (p.d.f) of data, which is usually required in simulations. However, if we generate samples with the same size as the original sample, we will be nearly certain to generate samples that include certain items of the original sample more than once. In several situations, this method may cause serious problems because the influence of the repeated items may be over magnified.

To overcome this shortcoming, we utilized the smoothed bootstrapping approach recommended by Simar and Wilson (1998) when we bootstrapped the efficiency measurement. This approach combines the advantages of simulation and bootstrapping. In smoothed bootstrapping, we need to define a p.d.f for the original sample. However, unlike in simulation, we do not need to estimate the exact p.d.f. We generate new samples through naïve bootstrapping method and then perturb the new samples through a standard error σ of the defined p.d.f to obtain the "smoothed" sample.

3.1.4.1 Definition and estimation of the distribution function

As mentioned earlier, for smooth bootstrapping, a p.d.f for the production set must be defined, although an exact estimation is not necessary. Compared with the efficiency score and productivity index estimation, sample p.d.f can be also defined and estimated through the parametric and nonparametric approaches. The parametric approach assumes that the samples come from a standard form of p.d.f defined by several parameters (such as normal distribution function, which is defined by its first and second moments, for example). By contrast, the nonparametric approach does not assume any form of p.d.f for the sample, but the p.d.f must be defined and estimated at each discrete point of the sample. The most widely used nonparametric method for the p.d.f estimation is the kernel method, which can be expressed by the following:

$$\hat{f}(z) = n^{-1}h^{-1}\sum_{i=1}^{n} K\left(\frac{z-Z_i}{h}\right)$$

where $K()$ is the function form (also called kernel; usually we use normal distribution as the kernel) used to estimate the p.d.f near each sample point Z_i (local p.d.f). The kernel determines the shape of the curve of the local p.d.f. The estimated p.d.f $\hat{f}(z)$ is clearly the average of kernel $K()$ centered at each sample point Z_i. $h > 0$ is the smooth factor (also called window width) of the kernel estimator. h scales down the distance of each point z from sample point z_i and determines how rapidly the tail of the local p.d.f centered around z_i will decrease, thereby attuning the dispersion of the n local p.d.f. $K()$. The larger the h is, the smoother $\hat{f}(z)$ becomes. By contrast, the smaller the h is, the rougher $\hat{f}(z)$ it will look.

In this chapter, following Wilson (1998),for simplicity I do not directly bootstrap production set (x, y), instead, I bootstrap the distance measure D. Indirect sample (x, y) is easy to obtain from D. However, the kernel method described earlier is a biased and inconsistent estimator of the p.d.f. of the distance measure D. Because the p.d.f of input or hyperbolic-oriented distance $\hat{\gamma}$ by definition is continuous on $[1, \infty)$, however, in practice, when we use Equation 3.3 to calculate $\hat{\gamma}$, we obtain a large number of $\hat{\gamma} = 1$. In other words, the c.d.f $\hat{F}(\gamma)$ is clouded near $\hat{\gamma} = 1$. The more inputs and outputs we choose in the DEA model, the higher possibility that a $\hat{\gamma}$ of DMU will be one. Moreover, estimating $F(\gamma)$ from $\hat{F}(\gamma)$ around $\gamma = 1$ is difficult because input- or hyperbolic-oriented distance measurement γ is low bounded by 1 (Simar and Wilson, 1998). To understand this issue, note that the local kernel estimated around $\gamma = 1$ requires data on both side of $\gamma = 1$.

Using the reflection method recommended by Silverman (1986) to "reflect" $\hat{\gamma}$ across the vertical line $\hat{\gamma} = 1$ can help solve the problem. For each point $\hat{\gamma} \geq 1$, we should reflect it to $\hat{\gamma}$-2 and estimate the distribution $\hat{f}(\mathbf{z})$ over the resulting joint set of 2n points. Therefore, the distribution function that we will estimate should be as follows:

$$G_h(z)=\frac{1}{2nh}\sum_{i=1}^{n}[K\left(\frac{z-\hat{\gamma}_i}{h}\right)+\left(\frac{z-(\hat{\gamma}_i-2)_i}{h}\right)]$$

Define:

$$\hat{F}_{n,h}(\hat{\gamma}_i)=\begin{cases}2G_h(\hat{\gamma}_i) & \text{if } z\geq 1,\\ 0 & \text{otherwise}\end{cases}$$

Shuster (1985) proved that $\hat{F}_{n,h}(\hat{\gamma}_i)$ is a consistent estimator of $F(\gamma_i)$.

Two other things should be decided: the choice of kernel *K*() and window width *h*. In practice, the function form used as kernel is usually not important. Like most of other researchers, I use the standard normal distribution function as the kernel in this chapter. The choice of *h* has much more important influence on estimate results. In this chapter, I use an equation suggested by Silverman (1986) to decide the *h*: $h=(4/5N)^{1/6}$, where N is the number of the original sample.

The smooth bootstrapping data generator is as follows:

$$\gamma^{**}=\begin{cases}\gamma^{*}+h\boldsymbol{\varepsilon}^{*} & \text{if } \gamma^{*}+h\boldsymbol{\varepsilon}^{*}\geq 1,\\ 2-(\gamma^{*}+h\boldsymbol{\varepsilon}^{*}) & \text{if otherwise}\end{cases} \tag{3.6}$$

where γ^{*} is the naïve bootstrapping sample by randomly drawing with replacement from the original sample γ. $\boldsymbol{\varepsilon}^{*}$ is randomly drawn from the kernel function ***K***. here ***K*** is a standard normal function. Thus, $\boldsymbol{\varepsilon}^{*}$ is a white noise. It can be proved (Simar and Wilson, 1998) that $\gamma^{**} \sim \hat{F}_{n,h}(\hat{\gamma}_i)$ and:

$$E(\gamma^{**}/\hat{\gamma}_1,\cdots,\hat{\gamma}_n)=\overline{\hat{\gamma}}$$
$$VAR(\gamma^{**}/\hat{\gamma}_1,\cdots,\hat{\gamma}_n)=\sigma_{\hat{\gamma}}^2+h^2,$$

where $\overline{\hat{\gamma}}$ and $\sigma_{\hat{\gamma}}^2$ are the average and variance of the estimated original scores, respectively.

To let γ^{**} has the same asymptotic moments as the original sample, we need transform γ^{**} according to the following:

$$\gamma^{**'} = \overline{\gamma}^{*} + \frac{1}{\sqrt{1+h^2/\sigma_{\hat{\gamma}}^2}}(\gamma^{**} - \overline{\gamma}^{*}) \tag{3.7}$$

$\overline{\gamma}^{*}$ is average of the naïve bootstrapping sample. The following can be proved:

$$E(\gamma^{**}/\hat{\gamma}_1,\cdots,\hat{\gamma}_n) = \overline{\hat{\gamma}}$$

$$VAR(\gamma^{**}/\hat{\gamma}_1,\cdots,\hat{\gamma}_n) = \sigma_{\hat{\gamma}}^2\left(1+\frac{h^2}{n(\sigma_{\hat{\gamma}}^2+h^2,)}\right)$$

Therefore, asymptotically, γ^{**} has moments same as that of the original estimates.

3.1.4.2 Data generation process

In sum, the smooth bootstrapping process for the unit can be explained as follows:

Step 1: Calculate the efficiency measurement $\hat{\gamma}$ from the original sample by Equation (3.4) or (3.5).

Step 2: Form a bootstrapping sample matrix γ^{*} by randomly drawing with replacement from the original sample $\hat{\gamma}$.

Step 3: obtain γ^{**} according Equation 3.5.

Step 4: Transform γ^{**} to $\gamma^{**'}$ according to Equation 3.6.

Step5: After $\gamma^{**'}$ is obtained, generate the bootstrapping production set $Q^{*}(x_i^{*}, y_i^{*}), i = 1\cdots n$. For hyperbolic distance $\gamma, x_i^{*} = \gamma_i^{**'}\hat{\gamma}_i^{-1}x_i$ and $y_i^{*} = \gamma_i^{**'-1}\hat{\gamma}y_i$.

Step 6: Compute the hyperbolic-oriented distance score set $\hat{\gamma}_i^{*}$, $i = 1\cdots n$, for the production set $Q^{*}(x_i^{*}, y_i^{*})$.

Step 7: Repeat Steps 2 to 5 B times to obtain the bootstrapping set $\hat{\gamma}_{ij}^{*}$, $i = 1\cdots n$, $j = B$.

3.1.4.3 Confidence interval for $\hat{\gamma}$ and hypothesis testing with bootstrapping

Using $\hat{\gamma}_{ij}^{*}$, $i = 1 \cdots n, j = B$ generated from the smoothed bootstrapping process, it is easy to estimate the confidence interval for $\hat{\gamma}$ at each point i in the sample. Because the distribution of $\hat{\gamma}_{ij}^{*}$ is unknown, I use a method called quantile approach to estimate the confidence intervals. Arrange the B number of $\hat{\gamma}_{ij}^{*}$ from the lowest to the largest:

$$\hat{\gamma}_{i1}^{*} \le \hat{\gamma}_{i2}^{*} \le \cdots \le \hat{\gamma}_{iB}^{*}$$

According to the principle of order statistics, if $\hat{\gamma}_{ij}^{*}$ is an random i.i.d variable, with $C.D.F = F(x)$, and p.d.f $= f(x)$, then $\hat{\gamma}_{ij}^{*}$ is normally distributed with a mean x_p and a variance $p(1-p)/[nf^2(x)]$, that is,

$$\hat{\gamma}_{ij}^{*} \sim N\left[x_p, \frac{p(1-p)}{n[f(x)]^2}\right]$$

where $p = l/B$, l is the order of $\hat{\gamma}_{ij}^{*}$, which is the number of $\hat{\gamma}_{i}^{*}$ that is less than γ_{ij}^{*}. x_p is the corresponding x when $F(x) = p$, that is, $x_p = F^{-1}(p)$.

Therefore, if we want the low bound of $\hat{\gamma}_{i}^{*}$ with p confidence level, we need only find the $\hat{\gamma}_{ij}^{*}$ with order $l = N(1-p)$. Similarly, if we want the upper bound of $\hat{\gamma}_{i}^{*}$ with p confidence level, we need only find the $\hat{\gamma}_{ij}^{*}$ with order $l = Np$.[3]

---

[3]For example, if B=1000 and we want the lower bound of $\hat{\gamma}_{i}^{*}$ at 95% confidence level, we need only to find the $\hat{\gamma}_{ij}^{*}$ whose order is 50 because the proportion of $\hat{\gamma}_{i}^{*}$ that is smaller than $\hat{\gamma}_{i50}^{*}$ is 5%. Following the same reasoning, if we want the upper bound of $\hat{\gamma}_{i}^{*}$ at 95% confidence level, we need only to find the $\hat{\gamma}_{ij}^{*}$ whose order is 950 because the proportion of $\hat{\gamma}_{i}^{*}$ that is higher than $\hat{\gamma}_{i950}^{*}$ is also 5%.

In a similar way, we can test hypotheses by using bootstrapping results. For example, if we are interested in whether, on average, efficiency scores in one group ($\overline{\gamma}_1$) is significantly different from that in another group ($\overline{\gamma}_2$), then for each sample j generated from bootstrapping, we calculate its average $\overline{\gamma}_{tj}^* = \frac{1}{N}\sum_{i=1}^{N}\gamma_{ti}^*$, t=1,2, $i = 1,\cdots,N$; $j = 1,\cdots,$ B . Thereafter, we calculate the statistic $t_j = \overline{\gamma}_{2j}^* - \overline{\gamma}_{1j}^*$, $j = 1\cdots B$. By using quantile approach, we can calculate the left-side significance level (*SL*) of the hypothesis $t_j < 0$ as

$$SL_{t_j<0} = \#(t_j < 0)/B,\ j = 1\cdots B \tag{3.7}$$

where $\#(t_j < 0)$ is the number of t_j , which is less than 0. In a similar way, we can also obtain the right-side significance level ($t_j > 0$) .

3.2 Measurement and estimation of the productivity changes

3.2.1 Estimating the Malmquist Index and its components

Given a panel database, besides estimating efficiency of DMUs for a fixed year, we can measure productivity changes over time. Productivity is an absolute concept. It means the ability of a DMU to transform inputs into outputs. It not only reflects the efficiency of a DMU, but also reveals the production technology level (production potential) in a period. From a static point of view, the measurements of productivity and efficiency are nearly the same. However, from a dynamic point of view, these two measurements are slightly different. For over a given period, not only the efficiency of a DMU but also the technology of the whole industry (i.e., production potential) may be changed. Thus, compared with the static efficiency measurement, the measurement of productivity changes not only provides a dynamic view of the productivity but also offers information on the changes in industrial technology.

In case of multiple outputs and inputs, economic and management analysts have used various approaches to measure productivity changes. Roughly, these approaches can be divided into two groups. The first group uses some kind of profit measurement. In this kind of approach, price information for inputs and outputs is needed. Indexes belonging to this approach include Törnqvist, Fisher, Paasche, Laspeyres, and Bennet–Bowley index. The other kind is based on production technical analyses. Among these approaches, Malmquist Index is the most widely used. Malmquist Index was first suggested by Malmquist for the purpose of analyzing consumer behaviors. Caves, Christensen, et al. (1982) are the first to suggest the theoretical possibility to use Malmquist index to analyze the productivity changes.

The Malmquist index involves using Shephard distance as efficiency measure. In addition to calculating the distance of DMU i in year t according to the frontier of year t as $D_i^{t,t}$, we can estimate the distance of DMU i in year $t+1$ according to the frontier of year t as $D_i^{t,t+1}$. Similarly, we can calculate the distance of DMU i in year t and $t+1$ both according to the frontier of year $t+1$ as $D_i^{t+1,t}$ and $D_i^{t+1,t+1}$, respectively. The Malmquist index is the geometric average of the two ratios:

$$M_i = \left[\frac{D_i^{t,t+1}}{D_i^{t,t}} \times \frac{D_i^{t+1,t+1}}{D_i^{t+1,t}} \right]^{1/2} \tag{3.8}$$

The explanation of M depends on the method used to calculate distance D. For input- or hyperbolic-oriented distance, a larger (smaller) value of M means deterioration (improvement) of productivity over time, whereas for the output-oriented distance, a larger (smaller) value of M means an improvement (deterioration) in productivity over time.

The Shephard distance used in the Malmquist index can be estimated with the DEA approach. Thus, no information on input and output prices is required. Furthermore, the index can also be easily decomposed to analyze sources of productivity changes. In this dissertation, I use the Malmquist index to measure productivity changes.

The Malmquist index is explained in Figure 3.3.

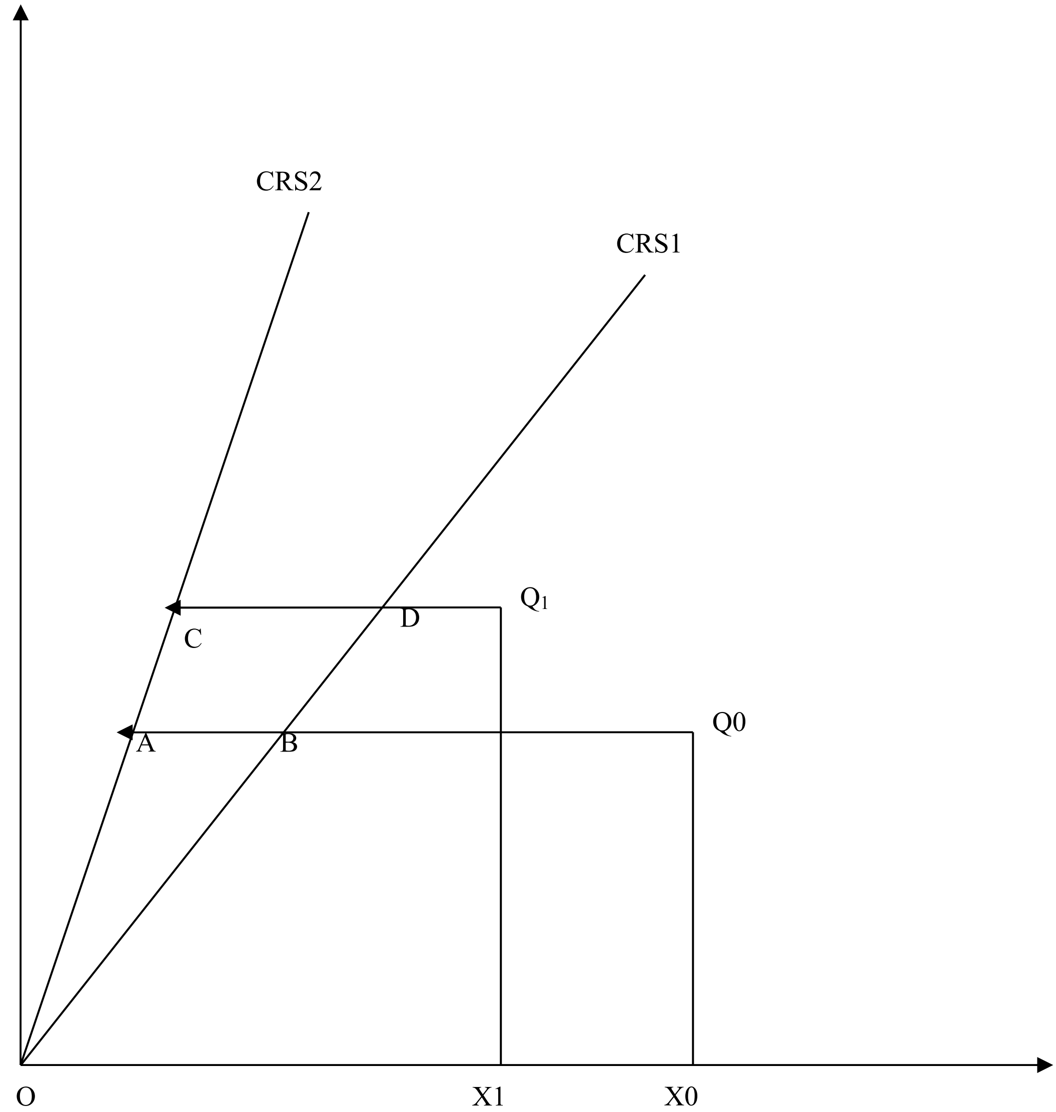


Figure 3.3: Malmquist index under the assumption of CRS

In Figure 3.3, the frontier in year *t* is CRS1. It moved to CRS2 in year *t+1*. DMU *i* is positioned in Q_1 and Q_2 in year *t* and *t+1*, respectively. Thus for DMU *i,* its input-oriented distance $D_i^{t,t} = OX_0 / BQ_0$, $D_i^{t,t+1} = OX_1 / DQ_1$, $D_i^{t+1,t} = OX_0 / AQ_0$,and $D_i^{t+1,t+1} = OX_1 / CQ_1$. Thus, the Malmquist index is as follows:

$$M = [(\frac{OX_1}{DQ_1}) / (\frac{OX_0}{BQ_0}) \times (\frac{OX_1}{CQ_1} / \frac{OX_0}{AQ_0})]^{\frac{1}{2}}$$

To detect sources of productivity changes, we must decompose the Malmquist index into several components. One of the most widely used decomposition methods was first proposed by Fare, Grosskopf, et al. (1992):

$$
\begin{aligned}
M &= \left[\frac{D_c^{t,t+1}}{D_c^{t,t}} \times \frac{D_c^{t+1,t+1}}{D_c^{t+1,t}} \right]^{1/2} \\
&= \frac{D_v^{t+1,t+1}}{D_v^{t,t}} \times \left[\left(\frac{D_c^{t+1,t+1}}{D_v^{t+1,t+1}} \right) \Big/ \left(\frac{D_c^{t,t}}{D_v^{t,t}} \right) \right] \times \left[\frac{D_c^{t,t+1}}{D_c^{t+1,t+1}} \times \frac{D_c^{t,t}}{D_c^{t+1,t}} \right]^{1/2} \\
&= EC \times SC \times TC
\end{aligned}
\qquad (3.9)
$$

where

$$
EC = \frac{D_v^{t+1,t+1}}{D_v^{t,t}};\ SC = \left(\frac{D_C^{t+1,t+1}}{D_v^{t+1,t+1}} \right) \Big/ \left(\frac{D_C^{t,t}}{D_v^{t,t}} \right) = \frac{SE^{t+1}}{SE^{t}}\ ;\ TC = \left[\frac{D_c^{t,t+1}}{D_c^{t+1,t+1}} \times \frac{D_c^{t,t}}{D_c^{t+1,t}} \right]^{1/2}
$$

D is the Shepherd distance defined as in Section 3.1.1, except that the low subscript V and C are distances calculated under VRS and CRS assumptions, respectively. EC is the change in pure efficiency. It is also called the "catch-up effects," because it measures the change in position of a given DMU to the current frontier between year *t* and *t+1* under VRS assumption. *SC* is the change in scale economy, and the ratios SE^t and SE^{t+1} are indices for scale economy in year *t* and *t+1*, respectively. *SE* is the ratio of CRS distance to VRS distance, which is the same as the *SC* defined in Equation 3-1. *TC* is the technological changes. It is the geometric average of two ratios. The first item is the distance ratio of a DMU *i* in time *t* according to the CRS frontier in *t* to the one in *t+1*. The second term is the distance ratio for the same DMU in time *t+1*. Given that *TC* measures productivity changes caused by the movement of the CRS production frontier from period *t* to *t+1*, it is also called as "frontier shift effects." Note unlike other components of the index, the distances in time *t* are in the numerator, whereas the distances in time *t+1* are in the denominator. Therefore, for hyperbolic-oriented distance, if *TC* less than 1(larger than 1), the frontier shifts outward (inward).

3.2.2 Bootstrapping the Malmquist index

Like in the case of efficiency estimation, Malmquist scores can be also bootstrapped. However, the Malmquist index is considerably complex because it involves efficiency scores in two periods.

3.2.2.1 Estimated distribution for Malmquist index

Given that the Malmquist index calculation involves efficiency distance measurement D_t in two periods, the calculation may encounter the problem of serial correlation, which often exists in time series data. That is, normally a DMU with high efficiency score in period *t* also has high efficiency score in period *t+1*. To overcome this problem, instead of estimating the univariate density function of D_t (the distance in time *t*), we should estimate the bivariate distribution of $\left[D_t, D_{t+1}\right]$. That is:

$$\hat{f}(\mathbf{z}) = n^{-1}h^{-2}\sum_{i=1}^{n} K\left(\frac{\mathbf{z}-\mathbf{Z}_i}{h}\right)$$

where $\mathbf{z}=(z_1, z_2)$, $\mathbf{Z_i} = [D_t, D_{t+1}]$.

As in the univariate case, we must also reflect the joint set of $\left[D_t, D_{t+1}\right]$. Only the reflection now becomes much more complex. Define a partitioned matrix:

$$\boldsymbol{\Delta} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ 2-\mathbf{A} & \mathbf{B} \\ 2-\mathbf{A} & 2-\mathbf{B} \\ \mathbf{A} & 2-\mathbf{B} \end{bmatrix} = \left[\delta_{ij}\right], \quad i = 1, \cdots, n, \; j=1,2$$

where $\mathbf{A} = \left[D_{i,t}\right]$, $\mathbf{B} = \left[D_{i,t+1}\right]$, $i = 1, \cdots, n$.

Let the estimated covariance matrix of [**A, B**] to be:

$$\hat{\Sigma}_1 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix}$$

It is easy to show that $\hat{\Sigma}$ is also the covariance matrix of [2-**A, 2-B**], and the covariance matrix of [2-**A, B**] and [**A,** 2-**B**] is:

$$\hat{\Sigma}_2 = \begin{bmatrix} \hat{\sigma}_1^2 & -\hat{\sigma}_{12} \\ -\hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix}$$

Thus, the corresponding partitioned covariance matrix of $\boldsymbol{\Delta}$ is:

$$\Sigma_\Delta = \begin{bmatrix} \hat{\Sigma}_1 \\ \hat{\Sigma}_2 \\ \hat{\Sigma}_1 \\ \hat{\Sigma}_2 \end{bmatrix} = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \\ \hat{\sigma}_1^2 & -\hat{\sigma}_{12} \\ -\hat{\sigma}_{12} & \hat{\sigma}_2^2 \\ \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \\ \hat{\sigma}_1^2 & -\hat{\sigma}_{12} \\ -\hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix}$$

Define a kernel bivariate density function:

$$\hat{G}_h(z) = \frac{1}{4nh^2}\sum_{j=1}^{4n} K_j\left(\frac{\mathbf{z}-\mathbf{\Delta}_j}{h}\right)$$

where $\mathbf{\Delta}_j$ is the jth row of the matrix $\mathbf{\Delta}$, $\mathbf{z}$ is a 2×1 vector, and K_j is a multivariate normal distribution (kernel) with an estimated covariance $\hat{\Omega}_j$:

$$\hat{\Omega}_j = \begin{cases} \hat{\Sigma}_1, & \text{for } j=1,\cdots,N;\ 2N+1,\cdots,3N; \\ \hat{\Sigma}_2, & \text{for } j=N+1,\cdots,2N;\ 3N+1,\cdots,4N; \end{cases}$$

Shuster (1985) proved that:

$$f_h^*(z) = \begin{cases} 4\hat{G}_h(\mathbf{z}), & \text{if } z_1 \geq 1, z_2 \geq 1 \\ = 0 & \text{if otherwise} \end{cases}$$

is a consistent estimate of the bivariate density function $\hat{f}(\mathbf{z})$.

The smooth bootstrapping data generator now is:

$$\Delta^{**} = \begin{cases} \Delta^* + h\boldsymbol{\varepsilon}^* & \text{if } \Delta^* + h\boldsymbol{\varepsilon}^* \geq 1, \\ 2-(\Delta^* + h\boldsymbol{\varepsilon}^*) & \text{otherwise} \end{cases} \tag{3.10}$$

where $\Delta^* = [\delta^*_{ij}]$ $i = 1,\cdots,n$, $j=1,2$ is the naïve bootstrapping sample matrix by randomly drawing with replacement from original matrix Δ . $\boldsymbol{\varepsilon}^*$ is the $N\times2$ matrix formed by white noise randomly drawn from the kernel function ***K***. It can be proved (Simar and Wilson, 1998) that $\Delta^{**} \sim f_h^*(z)$ and:

$$E(\Delta^{**}/\Delta) = \overline{\Delta}$$
$$VAR(\Delta^{**}/\Delta) = \mathbf{\Sigma}_\Delta / \mathbf{\Sigma}_K + h^2,$$

where $\overline{\Delta} = [\overline{\delta}_{ij}]$ is the sample mean of matrix Δ . $\mathbf{\Sigma}_K$ and $\mathbf{\Sigma}_\Delta = [\sigma^2_{\delta_{ij}}]$ are the covariance matrix for Kernel K and $\mathbf{\Delta}$, respectively.

To let Δ^{**} exhibit the same asymptotic moments as the original estimates, we must transform Δ^{**} according to:

$$\Delta^{***} = \overline{\Delta}^* + \frac{1}{\sqrt{1+h^2\mathbf{\Sigma}_K / \mathbf{\Sigma}_\Delta}}(\Delta^{**} - \overline{\Delta}^*) \qquad (3.11)$$

$\overline{\Delta}^* = [\overline{\delta}^*_{ij}]$ is the sample mean matrix of the naïve bootstrapping sample.

The following can be proved:

$$E(\Delta^{***} / \Delta) = \overline{\Delta}$$

$$VAR(\Delta^{***} / \Delta) = \sigma^2_\Delta \left(1 + \frac{h^2}{n(\mathbf{\Sigma}_\Delta / \mathbf{\Sigma}_K + h^2,)}\right)$$

Therefore, asymptotically Δ^{***} has the same moments as the original estimates.

If we scale the covariance matrix of K to $\mathbf{\Delta}$ so that $\mathbf{\Sigma}_K = \mathbf{\Sigma}_\Delta$, Equation 3.11 can be simplified to:

$$\Delta^{***'} = \overline{\Delta}^* + \frac{1}{\sqrt{1+h^2}}(\Delta^{**} - \overline{\Delta}^*) \qquad (3.12)$$

3.2.2.2 Data generating process

For the multivariate case [**A, B**], the process can be explained as follows:

Step 1: Calculate the efficiency measurement $\left[D_t, D_{t+1}\right]$ from the original sample and form $\mathbf{\Delta}$;

Step 2: Form a bootstrapping sample matrix $\mathbf{\Delta}^*$ by randomly drawing with replacement from original sample $\mathbf{\Delta}$.

Step 3: Perturb $\mathbf{\Delta}^*$ by $h\varepsilon^*$ to obtain $\mathbf{\Delta}^{**}$ according to Equation 3.10.

Step 4: Transform Δ^{**} to $\Delta^{***'}$ according to Equation 3.12.

Step 5: After $\mathbf{\Delta}^{***'}$ is obtained, generate the bootstrapping production set $[Q^*_t(x^*_t, y_t),\ Q^*_{t+1}(x^*_{t+1}, y_{t+1})]$, where $x^*_t = D^*_t \hat{D}^{-1}_t x_t ,.\ y^*_t = D^*_t \hat{D}_t y_t$.

Step 6: Compute the Malmquist Index set $m^*_i,\ i = 1 \cdots n,$ from the production set $[Q_t^*(x_t^*, y_t),\ Q_{t+1}^*(x_{t+1}^*, y_{t+1})]$.

Step 7: Repeat the step 2 to 5 B times to obtain the bootstrapping set $m_{ij}^*,\ i = 1 \cdots n, j = B$.

As in the case of efficiency score, after the bootstrapping sample of Malmquist score and its components is obtained, we can establish confidence interval and run hypothesis testing for the scores.

For technical details of the smooth bootstrapping approach used in this research, see Simar and Wilson (1999).

Chapter 4 Robust estimation of efficiency and productivity changes in Japanese Shinkin banks

4.1 Introduction

In this chapter, I use the non-parametric approaches discussed in Chapter 3 to estimate the efficiency and productivity changes in the Japanese Shinkin banks from Fiscal Year (FY) 2001 to FY 2008. Shinkin banks are among the most important regional financial institutions in Japan. According to the "Shinkin bank act" published in June 1951, Shinkin banks are regional, non-profit and mutual financial institutions, aimed at servicing small and medium enterprises and local inhabitants.

At the beginning of the 21st century, the Japanese economy struggled out from a "credit crunch" caused by the burst of the "bubble economy." In the first half of the 2000s, Japan enjoyed a relatively high economic growth. This environment is favorable for the Shinkin banks to improve their productivities. The government financial support to the Shinkin banks began in 2004 (The "financial function strengthening act") was also helpful for Shinkin banks to improve efficiency. But this trend was broken in the latter half of the 2000s due to the deterioration of international economic environment caused by the U.S sub-prime debt crisis and sovereign debt crisis in Europe. After 2006, the economic growth trend turned head down (see Figure 4.1). The severe economic environment made it very difficult for Shinkin banks to maintain high productivities.

Unit: %

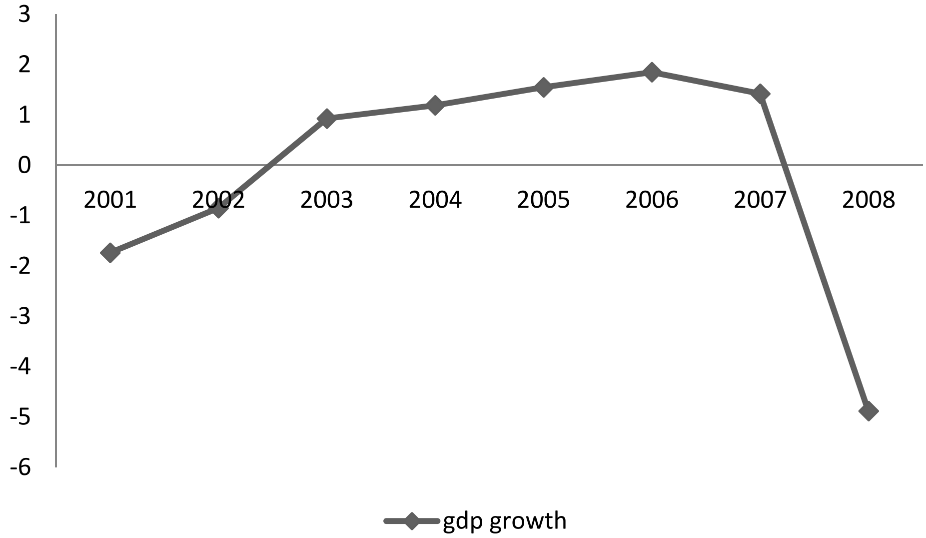


Figure 4. 1: The trend of GDP growth of Japan

Meanwhile, population structure of Japan has experienced essential changes. Because of the low birth rate, the population growth rate declined sharply from 2001 and it kept at negative levels since 2010 (see figure 4.2).

Unit: %

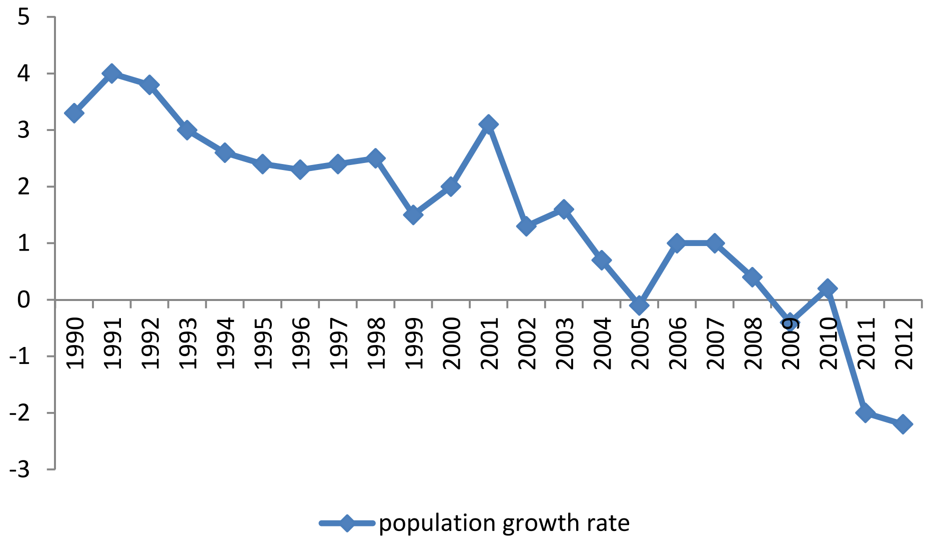


Figure 4. 2: The trend of population growth rate of Japan

The proportion of old people of the total population grew sharply. Japan gradually entered to the stage of aging society. This greatly increased the burden of the economy. In figure 4.3, we can see that since 1990, the aged dependence ratio (percentage of the population over 65 years old to the population aged from 15-64) increased rapidly.

Unit: %

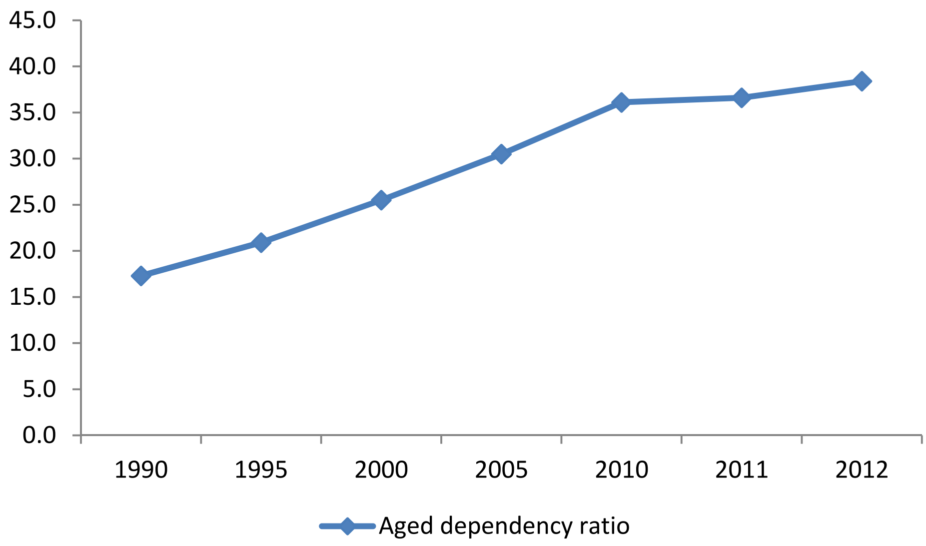


Figure 4. 3: The growth rate of aged dependence ratio of Japan

These changes in population structure were also supposed to be detrimental to the improvement of productivities in the Shinkin banks.

The environment of government supervision also changed. Many changes were introduced to Shinkin banks in this period. Their permitted scope of business was widened and many merger and acquisition（M&A）cases among Shinkin banks or Shinkin banks with other types of regional financial institutions occurred in the early 2000s. It is also hoped that these policy changes will have positive effects on the efficiency and productivity changes of the Shinkin banks.

All of these environment changes make Japanese Shinkin banks in this period an interesting case for the analysis of the sources of productivity changes in the regional banks. In this chapter I try to investigate how these environment changes have influenced the efficiency and productivity changes in Shinkin banks. I used an efficiency measurement called hyperbolic-oriented distance to estimate the efficiency. A nonparametric approach called data envelopment analysis (DEA) is used to estimate the efficiency scores. Based on these scores, an index called Malmquist index is used to measure the productivity changes. By analyzing estimated scores, we are able to examine the trends in efficiency and productivity changes in Shinkin banks during the sample period. The research further decomposes the estimated scores to inspect the sources of the trend. By comparing the efficiency and productivity changes between different sub-periods and among different groups, we can deduce the effects of environment changes on the efficiency and productivity changes in Shinkin banks as a whole and the differences of their effects upon different kinds of Shinkin banks.

The rest of this chapter is arranged as follows: Section 2 conducts a brief literature review on the empirical researches about the measurement and analysis of efficiency and productivity changes in financial institutions. Section 3 describes the data and the variables. Section 4 and 5 present and analyze the estimation results for efficiency and productivity changes, respectively. Section 6 draws conclusions from the analysis.

4.2 A short review about the literature of the empirical analysis of efficiency and productivity changes

Many papers have used the DEA approach to analyze the efficiency of financial institutions. For an early literature survey in this area, see Berger and Humphrey (1997). For a more recent survey, see Fethi and Pasiouras (2010). More recent papers include Devaney and Weber (2000) for the U.S. rural banking sector, Rebelo and Mendes (2000) for the Portuguese banking, Rezitis (2006) for Creek banking industry, Yang (2006) for the Canadian life and insurance companies, Garcia-cestona and Surroca (2008) for Spanish saving banks, Wheelock and Wilson (2008) for the U.S Federal Reserve check processing operations, Kao and Liu (2009) for the Taiwan commercial banks, Wheelock and Wilson (2009) for the U.S commercial banks. Following Wheelock and Wilson (2009), I use the hyperbolic-oriented distance as the measurement of efficiency. However, whereas they used the quantile estimation approach to deal with the stochastic characteristics of estimates, I use the smoothed bootstrapping method recommended by Simar and Wilson (1999) to set confidential intervals for the estimated scores.

There are also some papers analyzing the efficiency and productivity changes in Japanese financial institutions. Papers using the parametric approach include: McKillop et al. (1996), who analyzed the cost efficiencies of five giant Japanese city banks over the period 1978-1991 by estimating a composite cost function. They found on average there are statistically significant economy of scale among the sample banks, Altunbas et al. (2000), who investigated the impact of risk and quality factors on the pure and scale efficiency for a sample of 136 Japanese commercial banks (including city banks and regional banks) between 1993 and 1996. The innovation is they include in the model the ratio of loan loss provision to total loans as the indicator of quality of outputs and bank capital as well as ratio of liquid assets to total assets as the indicator of risk. They found that after considering the factors of quality and risks of the assets, the estimated economy of scale of the Japanese banks significantly declined, especially for those large banks. The above analysis all focused on the periods of financial crisis and restructuring. On the contrary, Tadesse (2006) analyzed the effects of consolidation on scale economy and technical changes by estimation a translog cost function. The sample includes both city banks and regional banks during the period of 1974-1991, which is a period of relative stability and high growth. He found there were diseconomies of scales among large banks; but he found the economy of scale in general increased over time and there were economy of scale among regional banks. Thus the analysis offers some rationale for M&A, especially among small and middle banks.

Meanwhile papers using DEA approaches include: Fukumaya (1993), who analyzed the technical and scale economy of 143 Japanese banks in FY 1991. The sample includes city banks and regional banks. He followed the intermediation approach. There were three inputs (labor, capital and funds from customers) and two outputs (revenue from loans and revenue from other business activities). He showed that the major cause of inefficiency was the lack of pure efficiency, not diseconomy of scale. He examined the relationship between the size of bank assets or bank income and economy of scale and found that there was a significantly positive correlation between bank size and economy of scale. Fukumaya (1996) also analyzed the technical efficiency and return to scale of credit unions in FY 1992 using both input and output-oriented model. His sample consisted of 435 active credit unions in Japan. He employed labor, capital and deposit as inputs and choose loan and securities as outputs. Therefore what he used was intermediary approach. He found there were considerable inefficiencies among credit unions in Japan. The major reason of inefficiency is the pure inefficiency, not lack of scale economy. The larger the scales of the credit unions were, the higher their pure and total efficiency were. Thus he supports M&A among large credit unions or credit unions with commercial banks, since it will increase the scale of the merged credit unions. Fukumaya and Weber (2008) continued the study for the period of FY 2002-FY 2004. The number of samples varies from 118 to 126 for different sample years and includes both city and Shinkin banks. The DEA model they used is a directional output distance model. They used labor, physical capital and raised funds as inputs, and loan, other interest bearing assets and non-performing loans as outputs. Their specification is consistent with intermediation approaches. However, the innovation of their research is they used NPL as an undesired (bad) by-product in the credit production. They focused on the estimation of the shadow prices of the NPLs. Their conclusion is researchers should control for effects of NPLs in their estimation of efficiency of financial institutions. Drake and Hall (2003) estimated the efficiency of Japanese banks in FY 1996 for a sample of 149 banks by using a input-oriented DEA model. The sample include all the "ordinary banks" (city banks and regional banks) operating in Japan. He followed the intermediation approach. Their choice of inputs consists of three inputs: general and administrative expense, fixed assets and retail and wholesale deposits. The outputs also were three: total loan and bills discounted, liquidity assets and other investments in securities. He found that the sample banks exhibited substantial total inefficiency and the major cause of it was pure inefficiency, not economy of scales. He also found that large city banks have the least potential for improvement of pure efficiency or economy of scale. Thus encouraging M&A among large banks will not significantly improve the efficiency of the banking industry. On the other hand, the smallest banks have the lowest level of economy of scale. Thus they support M&A among the small banks. They also found powerful positive relationship between bank size and efficiency. Drake et al. (2009) further used a slack based model (SBM) to estimate the efficiency of Japanese financial institutions during the period of 1995-2002 and compared the efficiency scores under different approaches. For estimation using intermediation approaches, the average of the efficiency scores is high and dispersion of the scores is low. This means there is no much room for improving the efficiency of the banks. Thus encouraging M&A among banks and other restructuring policies will not significantly improve the efficiency of the large banks. On the other hand, they found that estimates using production approaches have much lower average scores of efficiencies and the differences of the scores among banks are very large. Thus there are rooms to improve the efficiencies of the banking industry as a whole by using M&A and other restructuring policies. Horie (2010) analyzed the relationship between the productivity changes and operating areas in the Japanese Shinkin banks during the period of FY 2005-FY2007 by an input-oriented model. The sample included 257 Shinkin banks which were in continuous operation during the sample years. He divided the sample into four groups (large city, medium city A, medium city B and small city) and estimated a production frontier separately for each group. He used the production approach. There were two inputs (labor expenses and fixed expenses) and two outputs (interest income and other interest income) in his model. He found that there were considerable inefficiencies in all groups, but the Shinkin banks in the medium cities (normally with medium scale) were the least efficient. For all groups, the number of banks which were in IRS increased while the number of banks in DRS decreased during the sample period. This trend was more significant for the large and medium city group. He also estimated the Malmquist index of the sample banks to check the productivity changes in Shinkin banks during the period FY 2005-2007. He found that except for banks in large cities, on average the productivity of the Shinkin banks has notably declined and the larger city group Shinkin banks belong to, the better they performed in improving productivities. He decomposed the index into pure efficiency changes (catch up effect) and technical efficiency changes (frontier shifts) to explore the sources of the trend. He found that in improving pure efficiency, the small Shinkin banks perform best and was the only group which had improved their pure efficiency. This indicates Shinkin banks in small cities worked hardest in improving pure efficiency. The pure efficiencies in other groups all have declined. The Shinkin banks in medium city B group rank first in pure efficiency decline. On the other hand, for improving technology (frontier shifts effect), Shinkin banks in large cities had done best. Their technology had actually improved. This reflects the fact that banks in large cities are more able to employ new technologies (because they normally have large scale) and their economic environments were more favorable. The technical efficiencies of the Shinkin banks in small cities declined most fast. This maybe because banks in small cities were small; thus they were unable to follow the newest technologies. Also it may reflect the fact that the economic environments in small cities worsened most severely.

Among them Fukuyama (1996), Fukuyama and Weber (2008) and Horie (2010) are most closely related to the present research. Fukuyama (1996), Fukuyama and Weber (2008) emphasized the importance of the choice of direction of measurement. We follow the method used by Fukuyama and Weber (2008), but we use hyperbolic-oriented rather than directional distance as the measure of efficiency. Our choice of input and output is the same as Horie (2010). However, Horie (2010) divide the total sample into three sub-groups according to the type of cities they belong to and estimated a frontier for each group. In my case, I pool the sample together to estimate a single frontier. I do so for consideration of further regression analysis on the determination of efficiency and productivity changes in the next chapter, in which "operation areas" is one explanatory variable. I also group the sample into subgroups and check the differences of the estimates between different groups. But the division standards are little different. In the efficiency estimation, I group the sample according to the asset scales of the banks. In the estimation of productivity changes, I divided the sample according to their market shares. Another difference is that Horie's is deterministic in nature, mine is stochastic.

The analysis included in this chapter makes several contributions to the literature. First, I use the nonparametric bootstrapping approach suggested by Simar and Wilson (1999) to get robust estimates of the efficiency scores and Malmquist index. To my knowledge, this is the first attempt of such kind in the researches of Japanese financial institutions. Second, I use the hyperbolical-oriented distance instead of input- or output-oriented Shephard distance as the measure of efficiency. This is also the first attempt of using this measurement for the case of Japanese financial institutions. Like the directional distance measurement which was used by Fukuyama and Weber (2008), hyperbolical-oriented distance considers both output and input efficiency. With this measure, we avoid the problem of possible discrepancies between input- and output-oriented distances. Unlike directional distance, it is also easy to be decomposed. In addition, hyperbolic-oriented distance is closely related to the concept of profit, which is the conventional measure of efficiency.

4.3 The data

Our data consist of annual data from the income statements of the Shinkin banks from FY 2001 to FY 2008. The data is obtained from the database of Nikkei NEED[4].

In the estimation of efficiency and productivity changes, one difficult problem is the choice of time length. If the time length is too short, it is very likely that no significant changes in productivity will be detected. Besides, normally more than 2 years is needed for the effects of M&A to be fully exposed (Horie, 2010). For this reason, as well as for purpose of balance, I choose a 3 year time length. This divides the entire study period into two 3 year periods (i.e., FY 2001 to FY 2004 and FY 2005 to FY2008) and results in two estimations of the Malmquist indexes.

---

[4] This database is offered by the Company "Japanese Economic News" (Nihon keizai Shimbon, Nikkei.) The database includes various kinds of financial and economic data. FY 2001 and FY 2008 are the beginning and end year of the database when the paper is written.

M&A and shutdown may cause discontinuity in the data for acquired or closed banks. In the case of the acquiring banks or the merging banks (some with a new name), the operating environments are also significantly changed, making simple comparison of these banks before and after the merging misleading. To avoid this problem, for each period I excluded all of those banks which have been involved in the M&A activities or have been closed down during the period. After doing so, for the period FY 2001-FY 2004, the sample is reduced from 303 to 232. However, the sample for the period FY 2005-FY 2008 is rebound to 261 because of relatively few cases of M & A incidents in the later part of 2000s.

As mentioned in Chapter 3, one of the serious problems of the DEA is its high sensitivity to outliers. Therefore, to get robust estimates, we need a technique to detect and delete outliers from samples. However, most outlier detection techniques are designed for parametric methods. Here I use the approach suggested by Wilson (1993), which is particularly designed for nonparametric frontier models[5]. By using this technique, 6 outliers are detected in the first period so the sample is further reduced to 226. In the second period, 5 outliers are detected, reducing the sample to 256.

The choice of output and input is another important, but difficult task in DEA， especially in the case of financial institutions. There are two different definitions of financial institutions. The production approach treats financial institutions as organizations producing financial services. Meanwhile, the intermediation approach looks upon the financial institutions as a medium between debtors and creditors. That is: banks buy funds from some customers and sell assets to other customers. A major difference between these two definitions of financial institutions lies in input selection. In the production approach, only direct physical input, such as employees and operational spaces are treated as input. Deposits are considered as products offered to customers. On the other hand, in the intermediation approach, medium outputs, such as deposits, are considered as input. For a detailed description of the strengths and weaknesses of these two controversial definitions and their effects on the estimation

[5] The technique details of the method are not given in the paper due to the limitation of space, interested authors can refer to the paper by Wilson (1993).

of the efficiency of the financial institutions, see Berger and Humphrey (1997).

As for the measurement of the scale of output and input, both quantities and values (income or net income for output and cost for input) may be used. There are too many different kinds of financial services that a financial institution may offer, and directly pooling different kinds of products up to a few major categories is usually not possible, so quantity is not an ideal measure of production scale for financial institutions. Meanwhile, in using value measurement, we should keep in mind that values may reflect pricing ability instead of productivity levels.

As in Horie (2010), the method I used is similar to the production approach. I use value rather than the volume of output and input as the measures of scales. Since the scope of business of Shinkin banks is not as wide as that of large financial institutions in Japan, I focus on credit services provided by Shinkin banks, which accounts for more than 70% of the current incomes of most Shinkin banks. In the income statement of Shinkin banks, the credit activities are reflected under the entry of "Income on funds managed." I group items under this entry of income to form two products: A single item in the entry called "Interests from loans," which is the interest incomes from loans, forms the first product. It is the largest source of interest income of Shinkin banks. Meanwhile, other items in the entry, such as interest incomes from call loans, bonds， and deposits in other financial institutions, are aggregated to form the second product called "other interest income."

Unlike most analysts, I use net income rather than total income as output. That is, I deduct expenses on raising funds for a given credit from income gained from such credit. In this way, we not only can reduce one input in the model but also can avoid the difficult problem of treating deposits in the model. Interest earned from deposits is treated as income, whereas interest paid to depositors is treated as expenses incurred in the production of credit products

Unfortunately, there are no separate entries of interest expenses for each of the two products. All interest expenses are aggregated under a single entry "Fund Raising Expenses." To get corresponding expense for each of the two products, following Horie (2010), I divide this single entry into two entries by the weight of each product on the total interest incomes; thus the equations for the two products are as follows:

$$NY_L = Y_L - \frac{Y_L}{Y_I} C_I \tag{6}$$

$$NY_{NL} = Y_{NL} - \frac{Y_{NL}}{Y_I} C_I \tag{7}$$

where:

NY_L =net interest from loans;

NY_{NL} = net other interest income;

Y_L = total interest income from loans;

Y_{NL} = total other interest income;

$Y_I = Y_L + Y_{NL}$ = total interest income; and

C_I = total fund raising expenses.

In the input side, also two inputs are selected: One is the labor expenses. In the income statement of Shinkin banks, these expenses are recorded under the entry "Labor expenses"; however, this entry only includes the expenses on the formal employees. In recent years, like in other Japanese corporations, informal workers have accounted for an increasingly large proportion of the employees in the Shinkin banks. Expenses on these employees are included in the entry called "General expenses," and they account for about one third of this entry (Horie 2010). Due to the lack of information, we cannot segregate expenses on informal employees from general expenses and add them to labor expenses. Thus we should keep in mind that labor expenses do not include all of the cost of labor inputs for Shinkin banks.

Another input used in this research is fixed expenses, which roughly correspond to capital input for Shinkin banks. This input is the combination of two expense entries in the income statement of Shinkin banks: "General expenses" and "expenses on service transactions." General expenses include rents for stores, depreciation, expenses on advertisements, deposit insurance fees, outsourcing expenses, and expense on informal employees, among others. Expenses on service transactions include expenses on financial services by the Shinkin banks for their financial activities. This entry is neither large enough to be considered as a separate input nor too small to be ignored. Since these expenses are similar to some of the general expenses (e.g., outsourcing expenses and expenses on informal employees), I added them to general expenses.

The calculations involve data comparison across time, making the inflation effect a necessary concern. To eliminate this effect, we use the GDP deflator to deflate the data separately for each period, with the beginning year of each period as 100.

Table 4.1 summarizes the descriptive statistics of the variables used in this chapter for FY 2001 - FY 2004 and FY 2005 - FY 2008, respectively:

Table 4.1: Descriptive statistics for inputs and outputs

Unit：million yen

	y11	y12	x11	x12	y21	y22	x21	x22
FY2001 - FY 2004								
Min.	220	77	220	132	202	66	204	125
Median	2247	746.5	1658	1042	2093	618.5	1470	957.5
Mean	3643	1225	2423	1599	3521	1087	2196	1561
Max.	21010	9418	11970	9678	33100	9604	18990	13580
SD	3762	1337	2295	1617	4090	1294	2305	1737
FY 2005 - FY 2008								
Min.	397	63	354	202	287	136	303	212
Median	2902	998.5	1807	1314	2528	1079	1827	1297
Mean	4808	1475	2873	2161	4318	1666	2835	2178
Max.	31510	11880	18280	14030	26390	11050	17980	14220
SD	5478	1589	3001	2401	4936	1723	2948	2457

Note:

y11=Net interest from loans in the beginning year
y21=Net interest from loans in the end year
y12= Net other interest income in the beginning year
y22= Net other interest income in the end year
x11= Labor expenses in the beginning year
x21= General expenses in the end year
x12= Labor expenses in the beginning year
x22= General expenses in the end year

Min, Median, Mean, Max and SD are the minimum, median, mean, max, and standard error of the sample, respectively.

4.4 Results for the estimation of efficiency

In this section I estimate the hyperbolic-oriented efficiency of the Shinkin banks in FY 2005 and FY 2008. I use a package of the software R called FEAR to estimate the data. It was designed by P. W. Wilson (2008) particularly for the purposes of DEA. I first analyze the results for the total sample and then decompose them according to the levels of asset to explore the relationship between asset size and efficiency.

4.4.1 Results for the total sample

At first we examine the descriptive statistics for estimations for FY 2005 and FY 2008. These are outlined in Table 4.2.

Table 4.2: Descriptive statistics for the efficiency estimates

	crs		vrs		sc	
	FY2005	FY2008	FY2005	FY2008	FY2005	FY2008
Median	1.163	1.163	1.125	1.136	1.023	1.006
Mean	1.158	1.156	1.120	1.130	1.035	1.024
Max.	1.324	1.345	1.294	1.340	1.197	1.256
SD	0.068	0.069	0.072	0.076	0.039	0.040

Note: crs and vrs are the efficiencies estimated under the assumption of CRS and VRS, respectively. sc is the scale economy.

Because hyperbolical-oriented efficiency scores are low-bounded by 1, I do not list the minimum of the statistics. From Table 4.2, we can see that, in both years, there are significant inefficiencies in Shinkin banks, whether measured under the assumption of CRS or VRS. However, inefficiency in scale economy is much milder. From a dynamic point of view, on average the efficiency measured under assumption of VRS declined from FY2005 to FY2008, but there is considerable improvement in scale economy. This leads to no substantial change in efficiency measured under the assumption of CRS during the same period.

Using the bootstrapping hypothesis testing method described in Chapter 3 (equation 3.7), we can see that both the hypothesis that on average the efficiency scores measured under the assumption of CRS is lower in FY 2008 than in FY 2005 ($\overline{crs}^{*}_{08} - \overline{crs}^{*}_{05} < 0$) and the hypothesis that on average the scores of scale economy is lower in FY 2008 than in FY 2005 ($\overline{sc}^{*}_{08} - \overline{sc}^{*}_{05} < 0$) cannot be rejected with 100 significance level. On the contrary, the hypothesis that on average the efficiency scores measured under the assumption of VRS are higher in FY 2008 than in FY 2005 ($\overline{vrs}^{*}_{08} - \overline{vrs}^{*}_{05} > 0$) cannot be rejected with 94.05% significance level. The test results are listed in the fourth column of Table 4.3.

Table 4.3: Results of hypothesis tests for efficiency scores

Test	definition	hypothesis	significance			
			total	1st	2nd	3rd
d_{crs}	Average *crs* in FY 2008 is larger than that in FY 2005	$\overline{crs}^{*}_{08} - \overline{crs}^{*}_{05} > 0$	0	99	24.6	12.2
d_{vrs}	Average *vrs* in FY 2008 is larger than that in FY2005	$\overline{vrs}^{*}_{08} - \overline{vrs}^{*}_{05} > 0$	94.05	98.05	72.65	79.7
d_{sc}	Average sc in FY 2008 is larger than that in FY 2005	$\overline{sc}^{*}_{08} - \overline{sc}^{*}_{05} > 0$	0	68.95	1.2	9.95

To further examine the trend of efficiency of Shinkin banks from FY 2005 to FY 2008, next we examine the number of banks which are efficient. These data are listed in Table 4.4:

Table 4.4: Number of efficient banks

	crs		vrs		sc	
	core	up95%	score	up95%	score	up95%
FY2005	6 （2.34%）	8 (3.13%)	27 (10.55%)	28 (10.94%)	6 (2.34%)	8 (3.13%)
FY2008	8 （3.13%）	10 (3.91%)	22 (8.59%)	26 (10.16%)	8 (3.13%)	10 (3.91%)
difference	2 （0.78%）	2 (0.78%)	-5 (-1.95%)	-2 (-0.78%)	2 (0.78%)	2 (0.78%)

Note: Score is the original results of the estimation

Up95% is the upper bound at 95% significance level of the bootstrapping confidence interval

Data in bracket are percentages of the total sample

In FY 2005, 6 Shinkin banks (2.34% of the total sample) were efficient measured under the assumption of CRS. 8 banks (about 3.13% of the total sample) are robustly efficient with 95% significance. Comparatively, in FY2008, 8 Shinkin banks (about 3.13% of the total sample) were efficient. 10 banks (about 3.91% of the total sample) can be robustly assured with 95% significance that they are efficient. Both increased by 2 banks （0.78% of the total sample）compared to FY 2004. For the efficiency measured under the assumption of VRS, in FY 2004, 27 Shinkin banks (10.55% of the total sample) were efficient. 28 banks (about 10.94% of the total sample) are robustly efficient with 95% significance. In FY2008, 22 Shinkin banks (about 8.59% of the total sample) were efficient. 26 banks (about 10.16% of the total sample) can be robustly assured with 95% significance that they are efficient. Compared with FY 2005, it decreased by 5 (1.95% of the total sample) and 2 banks (0.78% of the total sample), respectively. All those banks which are efficient under the assumption of CRS are also efficient under the assumption of VRS. Consequently, the number of banks which are efficient in scale economy is the same as in *crs*. Thus, from examining the number of efficient banks, we can conclude that the number of efficient banks under the assumption of VRS (we can call it pure efficiency) has decreased, but the number of banks which is efficient in scale economy has increased. This made the number of banks which are efficient under the assumption of CRS slightly increased. The results confirm the conclusion we get from the analysis of descriptive statistics.

Therefore we can conclude that on average there was significant inefficiency measured under the assumption of CRS (total efficiency) during both sample years and the cause of it was the significant pure inefficiency (in efficiency measured under the assumption of VRS). On average, the efficiency under the assumption of VRS had significantly declined from FY2005 to FY 2008, but the scale economy has significantly increased in the same period. It overcame the negative influence of the *vrs* and let the efficiency measured under the assumption of CRS significantly increased. This reflects the facts that the economic environment during the latter half of the 21 century worsened, but Shinkin banks have tried to improve their efficiency by increasing their economy of scale through M&A.

The results are consistent with the finding of Horie (2010) for Shinkin banks from FY2001 to 2007; but our results are more robust. Thus they gave stronger supports for the conclusions. They are also in accord with the finding of Fukumaya (1993, 1996), Drake and Hall (2003), Drake et al for other kinds of financial institutions and during different periods. This proves that the trend of efficiencies found in this section was wide - reaching and lasted for a long time.

Using equation 3.2 in Chapter 3, we find that all those banks which were not efficient under the assumption of CRS were at DRS. This means most of Shinkin banks are oversized.

4.4.2 Results for the sub-groups

For a more detailed examination of the trend in efficiency of Shinkin banks during the 2000s, I further divide the total sample into subgroups according to the scale of the Shinkin banks. I divide the sample into three groups. Banks belonging to the first and fourth quantile of the total income of Shinkin banks compose the first and the third group respectively. Those between them (the second and third qunatile) form the second group. Table 4.5 shows the descriptive statistics of the efficiency results for the subgroups:

Table 4.5: descriptive statistics of the efficiency results for subgroups

	crs		vrs		sc	
	FY2005	FY2008	FY2005	FY2008	FY2005	FY2008
Median						
1st group	1.162	1.171	1.114	1.132	1.033	1.025
2nd group	1.155	1.154	1.146	1.152	1.006	1.002
3rd	1.175	1.162	1.083	1.108	1.058	1.022
Mean						
1st group	1.157	1.175	1.113	1.133	1.04	1.038
2nd group	1.156	1.149	1.139	1.143	1.015	1.005
3rd	1.161	1.15	1.085	1.1	1.071	1.048
Max.						
1st group	1.283	1.328	1.246	1.328	1.16	1.151
2nd group	1.324	1.345	1.294	1.34	1.078	1.036
3rd	1.305	1.26	1.218	1.23	1.197	1.256
sd						
1st group	0.065	0.068	0.066	0.076	0.031	0.043
2nd group	0.069	0.071	0.072	0.072	0.020	0.008
3rd	0.070	0.065	0.066	0.074	0.048	0.056

Note: see Table 4.2.

Table 4.5 shows that in both sample years, on average the second group is the lowest (most efficient) for efficiency measured under the assumption of CRS among the three groups. The trend of scale economy is similar as that of CRS. On the contrary, for efficiency measured under the assumption of VRS, in both sample years, the third group has the lowest average score (most efficient) among the three groups; while the second group has the highest average score (least efficient).

Using the bootstrapping hypothesis testing method[6], the hypothesis that, in FY 2005, the average score of *crs* of the second group is significantly larger than that of the first group($d_{crs21} = \overline{crs_2}^* - \overline{crs_1}^* > 0$) can be rejected at 21.3% significance level. But the hypothesis the average score of the third group is significantly larger than that of the second group ($d_{crs32} = \overline{crs_3}^* - \overline{crs_2}^* > 0$) cannot be rejected at 97.45% significance level. Therefore, we may conclude with confidence that in FY 2005 on average the third group has the lowest level of efficiency measured under the assumption of CRS.

[6] All the hypothesis test results in sub-group analysis are summarized in Table 4.6.

But there is no significant difference between the first and the second group. On the other hand, in FY 2008, the significance levels for the hypothesis d_{crs21} and d_{crs32} are 0 and 78.55%, respectively. The efficiency of the second group has relatively been improved compared to FY 2005 and it now was more efficient than the first group. Nonetheless there is no significant difference between the second and the third group. Thus we may conclude with confidence that in both sample years the second group remained the most efficient group under the assumption of CRS. The situation of scale economy is similar as the *crs*.

For efficiency measured under the assumption of VRS, in both sample years the hypothesis that the average score of the second group is significantly larger than that of the first group ($d_{vrs21} = \overline{vrs_2}^* - \overline{vrs_1}^* > 0$) cannot be rejected at varied significance level. But the hypothesis that the average score of the third group is significantly less than that of the second group ($d_{vrs32} = \overline{vrs_3}^* - \overline{vrs_2}^* < 0$) also cannot be rejected both at 100% significance level. Thus we can conclude that the efficiency measured under the assumption of VRS is lowest in the second group. The results are consistent with the finding from the original scores.

Table 4.6: Results of hypothesis tests for sub-groups

Unit: %

Test	definition	hypothesis	significance FY2005	significance FY2008
d_{crs21}	Average crs of group 2 is larger than that of group 1	$\overline{crs_2}^* - \overline{crs_1}^* > 0$	21.30	0.00
d_{crs32}	Average crs of group 3 is larger than that of group 2	$\overline{crs_3}^* - \overline{crs_2}^* > 0$	97.45	78.55
d_{vrs21}	Average vrs of group 2 is larger than that of group 1	$\overline{vrs_{12}}^* - \overline{vrs_1}^* > 0$	100	95.95
d_{vrs32}	Average vrs of group 2 is larger than that of group 1	$\overline{vrs_3}^* - \overline{vrs_2}^* > 0$	0.00	0.00
d_{sc21}	Average scale of group 2 is larger than that of group 1	$\overline{sc_2}^* - \overline{sc_1}^* > 0$	0.00	0.00
d_{sc32}	Average scale of group 2 is larger than that of group 1	$\overline{sc_3}^* - \overline{sc_2}^* > 0$	100	100

Dynamically, from Table 4.5, column 5-7, we see that, during the period of FY 2005 - FY 2008，for efficiency measured under the assumption of *CRS* and *VRS*, only the score of the first group on average has significantly increased (efficiency decreased). However, for scale economy, the average score of the second group has significantly decreased (efficiency increased). Therefore we may say that it is mainly due to the first group that the average efficiency measured under the assumption of VRS has declined during the sample period and the second group is the major cause of improvement in scale economy. The later mitigated the effects of the *vrs* on the *crs* and made the deterioration of *vrs* in the second and third group much less significant.

Second, we examine the number of banks which is efficient for each group. These data are listed in Table 4.7:

Table 4.7: Number of efficient banks for sub-groups

	CRS		VRS		SC	
	score	95%	score	95%	score	95%
1st						
FY2004	1 (1.56)	2 (3.13)	5 (7.81)	6 (9.38)	1 (1.56)	2 (3.13)
FY2008	1 (1.56)	1 (1.56)	2 (3.13)	3 (4.69)	1 (1.56)	1 (1.56)
difference	0 (0.00)	-1 (-1.56)	-3 (-4.69)	-3 (-4.69)	1 (0.00)	0 (-1.56)
2nd						
FY2004	4 (3.08)	5 (3.85)	9 (6.92)	9 (6.92)	4 (3.08)	5 (3.85)
FY2008	4 (3.13)	5 (3.91)	7 (5.47)	8 (6.25)	4 (3.13)	5 (3.91)
difference	0 (0.05)	0 (0.06)	-2 (-1.45)	-1 (-0.67)	0 (0.05)	0 (0.06)
3rd						
FY2004	0 (0.00)	1 (1.56)	2 (3.13)	2 (3.13)	0 (0.00)	1 (1.56)
FY2008	0 (0.00)	0 (0.00)	1 (1.56)	1 (1.56)	0 (0.00)	0 (0.00)
difference	0 (0.00)	-1 （-1.56）	-1 (-1.56)	-1 (-1.56)	0 (0.00)	-1 (-1.56)

Note: see table 3.4.

For efficiency measured under the assumption of CRS, in both sample years, the second group has the largest proportion of efficient banks. The third group gets the lowest proportion. Furthermore, compared to FY2005, in FY 2008 the proportion of efficient banks in the second groups have remained constant (for original scores) or slightly declined (for bootstrapping results); while the proportion of the efficient banks in the first and third group have decreased (only for bootstrapping results). But for efficiency measured under VRS, the situation is much complex, in FY2005, the first group has the largest proportion of efficient banks; Instead in FY 2008, the second group has the largest proportion of efficient banks. Still in both sample years the third group has the lowest proportion. Compared with FY2005, in FY 2008 in all groups the proportions of efficient banks have significantly declined. The lower the group, the higher is its decreasing rate. The finding is consistent with those from the original scores.

In conclusion， in both sample years, banks with medium scale of assets are least efficient measured under the assumption of VRS. But they are most efficient in scale economy. This compensated their poor achievements in efficiency measured under the assumption of VRS and made them also most efficient measured in the assumption of CRS. This is consistent with the theory of scale economy. From FY2005 to FY2008，the efficiencies of the banks with small scale of assets have declined. But the scale economies of the medium banks have increased.

The results are inconsistent with those of Fukumaya (1993) and Drake and Hall (2003) for city and regional banks, Fukumaya (1993) for credit unions. The division standard of Horie (2010) is a little different from that used here. But if we assume Shinkin banks located in large cities are normally also larger in scale, the sub-group results are roughly comparable. The findings in this section are roughly consistent with those of Horie (2010). But our findings are more robust.

4.5 Results for the estimation of productivity changes

In the above section we already contacted the concept of comparison of efficiency between different times. We now deal with this problem more formally. In this section, I estimate the productivity changes during the sample period by using an index called Malmquist index. By analyzing the productivity changes we not only can understand the changes of efficiency during the examined period, but also the technology changes during the period.

4.5.1 Results for the total sample

At first we examine the descriptive statistics for estimations on FY 2001-FY 2004 and FY 2005-FY 2008. These are outlined in Table 4.8.

Table 4.8: Descriptive statistics of the Malmquist results

	Malm		Pure.eff		Tech		Scale	
	FY01/04	FY05/08	FY01/04	FY05/08	FY01/04	FY05/08	FY01/04	FY05/08
Min.	0.7859	0.9509	0.8373	0.9066	0.973	0.9867	0.8933	0.8937
Median	1.006	1.034	1	1.001	1.002	1.038	1.004	0.9939
Mean	1.011	1.035	1.002	1.01	1.006	1.037	1.004	0.9896
Max.	1.317	1.217	1.149	1.206	1.157	1.09	1.074	1.068
SD	0.052	0.046	0.046	0.048	0.025	0.02	0.052	0.046

Note:

Malm = Malmquist index; Pure. eff = pure efficiency score;

Tech = technical efficiency score Scale = scale economy score

In both periods, the median is above 1, indicating less than half of the Shinkin banks have improved their productivity. The mean of the Malmquist index is also above 1. Using the bootstrapping method described in Chapter 3, we cannot reject the hypothesis that in both periods the average of Malmquist index ($\overline{malm}^*_{1j}$ and $\overline{malm}^*_{2j}$) are above 1 with 91.3% and 99.93 significance levels respectively[7]. Thus, we can conclude that in the first period, there is no significant decline in productivity; but in the second period, the productivity has significantly decreased. These results are

[7] All the hypothesis test results are summarized in Table 4.10

consistent with the findings of Horie (2010) for the period of FY 2005-2007.

All quantile statistics for the Malmquist index are higher in FY 2005-FY 2008 than in FY 2001 - FY 2004. At 100% significance level we find that on average the Malmquist index is higher in the second period than in the first period ($d_{malm} = \overline{malm}^*_{2j} - \overline{malm}^*_{1j} > 0,\ j = 1 \cdots B$) .

Thus we can conclude that from the first to the second period, productivity deterioration has worsened. However, the variance of the data (except for data on pure efficiency) are narrowed in the second period, indicating a tendency of convergence among Shinkin banks.

We then examine the number of banks whose productivity has improved (with Malmquist indexes less than 1) in each period. These data are listed in Table 4.9:

Table 4.9: Number of banks with improved Malmquist results

	Malm		Pure. eff		Tech		Scale	
	score	Up95%	score	Up95%	score	Up95%	score	Up95%
FY2001-FY2004	104	91	105	57	102	0	95	2
	(46.02)	(40.27)	(46.46)	(25.22)	(45.13)	(0.00)	(42.04)	(0.88)
FY2005-FY2008	61	35	110	68	9	0	178	34
	(23.83)	(13.67)	(42.97)	(26.56)	(3.52)	(0.00)	(69.53)	(13.28)
Difference	-43	-56	5	11	-93	0	83	32
	(-22.19)	(-26.59)	(-3.49)	(1.34)	(-41.62)	(0.00)	(27.50)	(12.40)

Note: Score is the original results of the estimation

Up95% is the upper bound at 95% confidence level of the bootstrapping confidence interval

Data in bracket are percentage of the total sample

In the period of FY 2001 - FY 2004, the productivity of 104 Shinkin banks (46% of the total sample) has increased. Among them 91 banks (about 40% of the total sample) can be robustly assured with 95% significance that their indexes are below 1. Comparatively，from FY2005 to FY2008, the Malmquist indexes of 61 Shinkin banks (about 24% of the total sample) were less than 1. Among them 35 banks (about 14% of the total sample) can be robustly assured with 95% significance that their index are below 1. Thus, from examining the number of banks whose productivity has been improved, we confirm the finding from the analysis of descriptive statistics.

Examining the components of the Malmquist index can shed light on the sources of the decline of productivity.

From Table 4.8, we see that in the first period the mean and median of the three components are almost the same (all slightly above 1). Using the bootstrapping hypothesis testing method, we find that, on average, none of the component scores are significantly different from the Malmquist scores. The significance levels for the hypothesis that, on average, pure efficiency and technical efficiency scores are higher than the Malmquist scores ($d_{ME1} = \overline{pure.eff}^{*}_{1j} - \overline{malm}^{*}_{1j} > 0$ and $d_{MT1} = \overline{tech}^{*}_{1j} - \overline{malm}^{*}_{1j} > 0$, respectively) are both 30.9%. In the case of scale economy, the significant level for ($d_{MS1} = \overline{scale}^{*}_{1j} - \overline{malm}^{*}_{1j} > 0$) is 38.5%.

From Table 4.9, we observe that in the first period, the number of banks with pure efficiency, technical efficiency, and scale efficiency scores less than 1 is also not much different from the number of banks with Malmquist indexes less than 1. Thus, from the estimates themselves, we find that the three components of the Malmquist index offer similar levels of contribution to the improvement of the index. However, when we consider the robustness of estimates, almost no Shinkin banks' technical or scale efficiency scores is robustly below 1.

Hence, we may conclude that in FY 2001-FY 2004, the major cause of the slight deterioration in the productivity of Shinkin banks is the worsening of technical and scale efficiency. This worsening of scale economy is somewhat surprising, because at in this period, the number of M&A cases is much larger than that in the later period. This is because all those banks which have engaged in M&A are deleted from the sample for the period the incidence occurred. Their effects will be reflected in the next period.

In the second period (FY 2005-FY 2008), the hypothesis that, on average, pure efficiency and scale economy scores are higher than the Malmquist scores ($d_{ME2} = \overline{pure.eff}^{*}_{2j} - \overline{malm}^{*}_{2j} > 0$ and $d_{MS2} = \overline{scale}^{*}_{2j} - \overline{malm}_{2j} > 0$) can be rejected at

99.85% and 100% significance levels, respectively. But the corresponding significance level for technical efficiency ($d_{MT2} = \overline{tech}^*_{2j} - \overline{malm}^*_{2j} > 0$) is 41.6%. It is not significantly different from Malmquist scores.

During the second period, the number of Shinkin banks with pure efficiency scores less than 1 is 110. Among them the number of banks whose upper bounds of 95% confidence intervals are below 1 is 68. Both are more than twice the corresponding level for the Malmquist index in the same period. For the scale efficiency component, 178 Shinkin banks have scores less than 1, the number is almost thrice the number of banks with Malmquist indexes less than 1, but only 34 banks have scores robustly below 1. On the other hand, only 9 banks have technical component less than 1. None of them is robust at 95% significance level. From these findings, we may conclude that in FY 2005-FY 2008, the major cause of the deterioration in the productivity changes of the Shinkin banks is the worsening of technical efficiency. Pure efficiency and scale efficiency play positive roles in the trend of productivity changes.

Comparing the two periods, the results of the bootstrapping hypothesis testing method show that the significance level is 84% for the statistic $d_{pure2} = \overline{pure}^*_{2j} - \overline{pure}^*_{1j} > 0$, 99.9% for the statistic $d_{tech2} = \overline{tech}^*_{2j} - \overline{tech}^*_{1j} > 0$ and 8.05% for the statistic $d_{scale} = \overline{scale}^*_{2j} - \overline{scale}^*_{1j} > 0$. Thus only the hypothesis that, on average, the trend of technical efficiency growth has been deteriorated in the latter period cannot be rejected with 95% significance level.

From Table 4.9, we see that, in FY 2005-FY 2008 to FY 2001- FY 2004, the proportion of Shinkin banks with pure efficiency scores less than 1 decreased by 3.49%. However, if considering the robustness of the results at 95% significance level, the ratio is increased by 1.34%. The proportion of banks with technical scores less than 1 has drastically dropped by 42%. However, as in the first period, none of the results are robust at 95% significance level. The ratio of Shinkin banks with scale efficiency scores less than 1 has noticeably increased by 27.5%. The ratio of banks with scores robustly less than 1 has increased by 12.4%.

Table 4.10: Results of hypothesis tests for total sample

Unit: %

statistic	definition	equation	significant
$\overline{malm}^*_{1j}$	Average Malmquist index in period 1is larger than 1	$\overline{malm}^*_{1j}>1$	91.3
$\overline{malm}^*_{2j}$	average Malmquist index in period 2 is large than 1	$\overline{malm}^*_{2j}>1$	99.93
d_{malm}	average Malmquist score in period 2 is large than in period 1	$\overline{malm}^*_{2j}-\overline{malm}^*_{1j}>0$	100
d_{ME1}	Average pure efficiency score is higher than total score in period 1	$\overline{pure.eff}^*_{1j}-\overline{malm}^*_{1j}>0$	30.9
d_{MT1}	Average technical score is higher than total score in period 1	$\overline{tech}^*_{1j}-\overline{malm}^*_{1j}>0$	30.9
d_{MS1}	Average scale economy score is higher than total score in period 1	$\overline{scale}^*_{1j}-\overline{malm}^*_{1j}>0$	38.5
d_{ME2}	Average pure efficiency score is higher than total score in period 2	$\overline{pure.eff}^*_{2j}-\overline{malm}^*_{2j}>0$	0.15
d_{MT2}	Average technical score is higher than total score in period 2	$\overline{tech}^*_{2j}-\overline{malm}^*_{2j}>0$	41.6
d_{MS2}	Average scale economy score is higher than total score in period 2	$\overline{scale}^*_{2j}-\overline{malm_{2j}}>0$	0
d_{pure2}	Average pure efficiency score is higher in period 2 than in period 1	$\overline{pure}^*_{2j}-\overline{pure}^*_{1j}>0$	84
d_{tech}	Average technical score is higher in period 2 than period 1	$\overline{tech}^*_{2j}-\overline{tech}^*_{1j}>0$	99.9
d_{scale}	Average scale economy score is higher in period 2 than in period1	$\overline{scale}^*_{2j}-\overline{scale}^*_{1j}>0$	8.05

Thus from the analysis of the components, we may also conclude that there is a significant deterioration in the productivities of Shinkin banks from the period of FY 2001-FY 2004 to the period of FY 2004-FY 2008. The major cause for this decline is the worsening of technical efficiency. However, we also see that scale efficiency has significantly improved, probably because the large number of M&A cases happened at the beginning of the century has gradually manifested its effects. Scale efficiency is the only component that has shown significant improvement.

4.5.2 Results for the sub-groups

For a more detailed examination of the trend in productivity changes of sinkin banks during the 2000s, as in the case of efficiency analysis, I further divide the total sample into subgroups to investigate results of the estimation. Instead of dividing the total sample according to the scale of shinklin banks, here I divide the sample according to an indicator reflecting the level of competition in areas where Shinkin banks operate. For regional financial institutions like Shinkin banks, regional environmental conditions are more important to productivity than scale. I am especially interested in the effects of M&A peaks at the beginning of the 21st century, so we only analyze results for FY 2005-FY 2008.

Because we measure Malmquist index in value terms，so the resulting scores reflect not only real productivity changes, but also the changes in price fixing ability of the banks. There are no clear theoretical assumptions on the influence of market power on productivity changes measured in value terms. Shinkin banks located in highly competitive areas may benefit from improvement in pure and technical efficiency because of intense competition (market competition hypothesis). On the other hand, Shinkin banks located in less competitive areas may benefit from high price fixing ability and improved economy of scale (market power hypothesis).

I used the indicator $share_i$, which is the ratio of the number of branches of Shinkin bank *i* to the total number of branches of all financial institutions operating in the same region, as the indicator of market power of the Shinkin banks.

I use “city, town or village” (shi, mura or machi) as the unit of area. Many Shinkin banks operate over more than one unit of area, so following Horie (2010), I use the weighted average of regional statistics of *share* as the indicator of market power:

$$Q_i = \sum_{j=1}^{M} w_{ij} share_{ij} \qquad i = 1, 2, \cdots N,$$

where：

Q_i is the weighted average of *share* for Shinkin bank *i*, in year *t* (for the reason of simplicity, *t* is omitted in the subscript);

w_{ij} is the weight given to area *j* for bank *i*, which is the ratio of the number of branches owned by bank *i* in area *j* to its total number of branches in year *t*; and

$share_{ij}$ is the market power indicator defined above of Shinkin bank *i* in area *j*.

We divide the sample into three groups. Banks belonging to the first and fourth quantile of the *share* distribution compose the first and the third group respectively. Those between them formed the second group. Table 4.11 shows the descriptive statistics of the Malmquist results for the subgroups:

Table 4.11：Descriptive statistics of the Malmquist results for sub-groups

	Malm	Pure.eff	Technical	Scale
Min				
1st Group	0.954	0.910	0.998	0.894
2nd Group	0.998	0.975	1.024	0.973
3rd Group	0.955	0.907	0.998	0.931
Median				
1st Group	1.036	1.000	1.028	0.996
2nd Group	1.033	1.004	1.039	0.994
3rd Group	1.034	0.999	1.049	0.991
Mean				
1st Group	1.037	1.014	1.028	0.995
2nd Group	1.035	1.011	1.038	0.987
3rd Group	1.034	1.003	1.043	0.989
Max				
1st Group	1.166	1.185	1.067	1.068
2nd Group	1.217	1.206	1.090	1.059
3rd Group	1.196	1.118	1.083	1.038
SD				
1st Group	0.047	0.047	0.018	0.027
2nd Group	0.046	0.050	0.019	0.027
3rd Group	0.043	0.044	0.020	0.023

Note: see Table 4.8.

Table 4.11 shows that from FY 2005 to FY 2008，on average, the higher the *share* group a Shinkin bank belongs to, the slower is its decline in productivity. Shinkin banks with larger market shares experience slower declines in productivities compared to the banks with smaller market shares. Note also that Shinkin banks in the first group vary more widely in productivity changes than other groups.

Using the bootstrapping hypothesis testing method[8] described in Chapter 3 (equation 3.7),, the hypothesis that, on average, the Malmquist index of the first group is significantly larger than that of the second group($d_{M21} = \overline{malm}^*_{22} - \overline{malm}^*_{21} > 0$) cannot be rejected at 99.65% significance level. But the hypothesis ($d_{M32} = \overline{malm}^*_{32} - \overline{malm}^*_{22} > 0$) can be rejected at 0.55% significance level. Therefore, we may conclude with confidence that the second group has the highest rate of decline in productivity. This is inconsistent with the inferences from the original scores.

Analyzing the descriptive statistics of the components can help us find the sources of the foregoing confusing results. We observe from Table 4.11 that in FY 2005 to FY 2008, for original results similar conclusions can be drawn for pure efficiency and scale economy. On the other hand, the trend of technical efficiency is different from the trends of the two other components. On average, the higher the share group a Shinkin bank belongs to, the faster is its decline in technical efficiency. This contradiction may have caused the confusing result of the robust analysis described above. These findings are in accord with the market competition theory.

Using the bootstrapping hypothesis testing method, the hypothesis that, on average, the pure efficiency score of the first group is significantly less than that of the second group ($d_{pure21} = \overline{pure}^*_{22} - \overline{pure}^*_{21} > 0$) cannot be rejected only at 51.25% significance level. Therefore the two groups are not significantly different from each other. On the other hand, the hypothesis that on average the pure efficiency score of the third group is larger than that of the second group ($d_{pure32} = \overline{pure}^*_{23} - \overline{pure}^*_{22} > 0$) cannot be rejected at 1.15% significance level. Similarly, the hypothesis $d_{tech21} = \overline{tech}^*_{22} - \overline{tech}^*_{21} > 0$ cannot be rejected only at 80.85% level; whereas, the hypothesis $d_{tech32} = \overline{tech}^*_{23} - \overline{tech}^*_{21} > 0$ cannot be rejected at 92.25% level. For scale

8 All the hypothesis test results in sub-group analysis are summarized in Table 4.13.

economy scores, the hypothesis $d_{scale21} = \overline{scale}^*_{22} - \overline{scale}^*_{21} > 0$ cannot be rejected only at 14.05% significance level; whereas, the hypothesis $d_{scale23} = \overline{scale}^*_{23} - \overline{scale}^*_{21} > 0$ cannot be rejected at 69.03% significance level. This means that, on average, the change in both technical and scale efficiency is not significantly different across the three groups.

Second, we examine the number of banks whose productivity has improved for each group. These data are listed in Table 4.12:

Table 4.12: Number of banks with improved Malmquist results for sub-groups

	Malm		Pure. Eff		Tehnical		Scale	
	score	Up95%	score	Up95%	score	Up95%	score	Up95%
1st	14 (21.88)	9 (14.06)	26 (40.63)	13 (20.31)	5 (7.81)	0 (0)	38 (59.38)	7 (10.94)
2st	33 (25.78)	16 (12.5)	52 (40.63)	36 (28.13)	3 (2.34)	0 (0)	95 (74.22)	19 (14.84)
3rd	14 (21.88)	10 (15.63)	32 (50)	19 (29.69)	1 (1.56)	0 (0)	45 (70.31)	8 (12.5)

Note: see Table 4.9

In FY 2005- FY 2008, the second group has the largest proportion of banks with improved productivities. However, after considering the robustness of the estimation, the third group now emerges with the largest proportion (this may partly explain the difference between the original scores and hypothesis testing). Pure efficiency and scale economy show a similar trend. Considering the robust results, now the first group has a significantly smaller proportion of banks with improved pure and scale efficiency than the other groups. For technical efficiency, the picture is different. Higher groups have smaller proportions of banks with improved productivity, and the differences are quite significant. However, since none of the results are statistically significant. The findings are without much meaning. These conclusions are consistent with inferences from the original scores.

Table 4.13: Results of hypothesis tests for sub-groups

Unit: %

Test	definition	equation	significance
d_{M21}	average Malmuist index of group 2 is larger than that of group 1	$\overline{malm}^{*}_{22} - \overline{malm}^{*}_{21} > 0$	99.65
d_{M32}	average Malmuist index of group 3 is larger than that in group 2	$\overline{malm}^{*}_{32} - \overline{malm}^{*}_{22} > 0$	0.55
d_{pure21}	average pure efficiency of group 2 is larger than that in group 1	$\overline{pure}^{*}_{22} - \overline{pure}^{*}_{21} > 0$	51.25
d_{pure32}	average pure efficiency of group 3 is larger than that of group 2	$\overline{pure.eff}^{*}_{1j} - \overline{malm}^{*}_{1j} > 0$	1.15
d_{tech21}	average technical score of group 2 is larger than that in group 1	$\overline{tech}^{*}_{22} - \overline{tech}^{*}_{21} > 0$	80.85
d_{tech32}	average technical score of group 3 is larger than that of group 2	$\overline{tech}^{*}_{23} - \overline{tech}^{*}_{21} > 0$	92.25
$d_{scale21}$	Average scale economy score of group 2 is larger than that of group 1	$\overline{scale}^{*}_{22} - \overline{scale}^{*}_{21} > 0$	14.05
$d_{scale32}$	average scale economy score in group 3 is larger than that in group 2	$\overline{scale}^{*}_{23} - \overline{scale}^{*}_{21}$	69.03

In summary, we can conclude that in FY 2005-FY 2008, Shinkin banks which located in least competitive areas have experienced the least decline in productivities. But this result is not robust. When we decompose the scores, banks located in least competitive areas also declined least in the pure efficiency. This result is robust. The results are in accord with the market power hypothesis. In contrast, for technical efficiency, banks located in more competitive areas perform better than those in less competitive areas, indicating that these banks benefit more from technological progress. But the result is not robust. This finding agrees with the theory of market competition. The contradiction between the two scores may have caused the ambiguity in the influence of market power on the productivity changes if we consider the robustness of the estimation. There are no significant differences in scale efficiency across the three groups.

Because the division standard is different, the results of Malmquist scores for sub-groups in this chapter are not totally comparable with those of Horie (2010). However, if we assumes the larger a city is, the more intense the competition in its financial market, then we can roughly compared the results of the two analysis. We find that for original scores of Malmquist score, the conclusion are completely different; however from the robust test we find that there are no statistically significant differences between different groups. The robust results for change of pure efficiency of this chapter are consistent with those of Horie (2010). But the robust results for technical changes are in contrast with those of Horie (2010).

4.6 Conclusion

In this chapter, I try to use a more robust approach than those used by other related papers to estimate the efficiency and productivity changes in Japanese Shinkin banks during the 2000s by using a bootstrapping technique. I find that on average the efficiency under the assumption of VRS had significantly declined from FY2005 to FY 2008, but the scale economy has significantly increased in the same period. It overcame the negative influence of the *vrs* and let the efficiency measured under the assumption of CRS significantly increased. In both sample years, banks with medium scale of assets are least efficient measured in the assumption of VRS. But they are most efficient in scale economy. This made them also most efficient measured in the assumption of CRS. This is consistent with the theory of scale economy. From FY2005 to FY2008，the efficiencies of the banks with small scale of assets have declined. But the scale economies of the medium banks have increased.

In FY 2001- FY 2004, there have been no significant changes in productivity. However, in FY 2005- FY 2008, productivity has significantly declined. The major source of this trend is the deterioration in technical efficiency (the inward shift of the production frontier). This finding is in accord with the deterioration of economic environment in the latter half of the 2000s. However, in the second period, the scale efficiency of Shinkin banks have notably improved, which partly offset the deterioration of the environments. This may have originated from the time lag in the effects of active M&A in the early 2000s.

Using the technique of bootstrapping greatly helps us to determine the sources of the trends of the productivities. For example, I find that, in the first period, a large proportion of Shinkin banks have improved their technical efficiency. Yet after considering the robustness of the results, we find the contribution of technical efficiency is drastically reduced.

Grouping banks in the total sample according to the level of competition reveals the relationship between market power and productivity changes. From the original scores, I find that banks located in the least competitive areas experienced the least declines in productivity, but this result is not robust. Checking the components, we observe that banks located in the least competitive areas experienced the slowest declines in pure efficiency. However, banks located in highly competitive areas are more successful in their efforts of slowing down the decline in technical efficiency. This compensates for their weakness in pure efficiency and makes the results less clear.

This chapter leaves some questions unanswered. It only analyzes the effects of competition on the production process. In next chapter, I will analyze the influence of other external factors which do not directly involved in the production process, but may also influence the productivity.

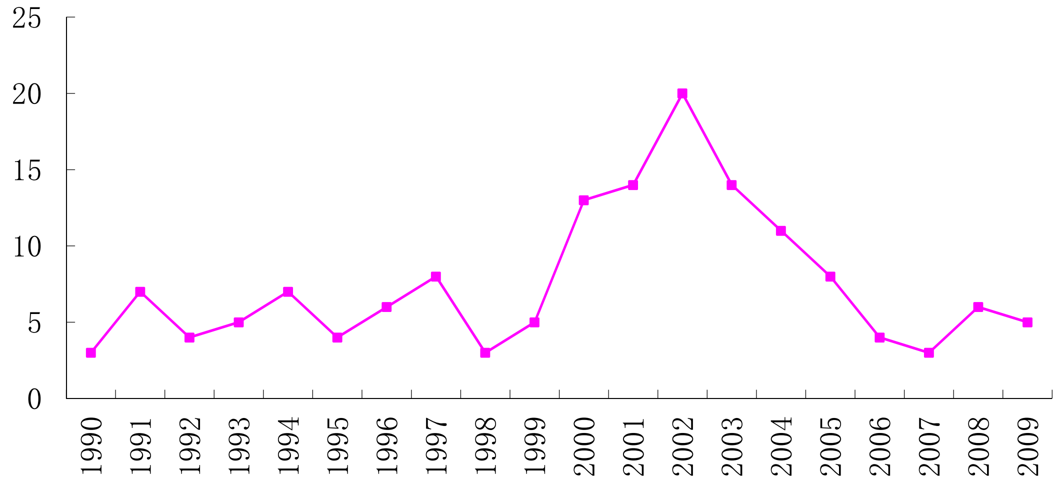
Chapter 5 The effects of M&A on the efficiency and change of productivities

—evidence from the Japanese Shinkin banks

5.1 Introduction

In Chapter 4, I estimated the efficiency and productivity changes in Shinkin banks using chosen inputs and outputs and analyzed their sources. In this chapter, I further explore the environment factors that determine the efficiency and productivity changes in Shinkin banks. I am especially interested in the effects of M&A incidents on the efficiency and productivity changes.

From the beginning of the 21st century, the supervision environment for financial institutions in Japan experienced great changes. With the progress of deregulation and financial liberalization, the permitted business scope of the financial institutions was widened and the boundary between different kinds of financial institutions became obscured. To strengthen the financial system, the Japanese government encouraged M&A activities between financial institutions. This triggered a wave of M&A among Shinkin banks at the beginning of 2000s (see Figure 4.1). As the result, the number of Shinkin banks deceased from 371 at the end of FY 2000 to 279 by the end of FY 2008. It is interesting to analyze the effects of this peak of M&A activities on the productivity changes in Shinkin banks.



Sources of data： Annual reports of Shinkin banks by Central Shinkin banks

Figure 5. 1: Number of M&A cases among Shinkin banks in 2000s

The influence of M&A is the key concern of many papers about the determination of efficiency and productivity changes of financial institutions. Interests in this area are not purely due to academic curiosity. They also come from policy considerations. Encouraging M&A among financial institutions is an important part of bank restructuring policy packages in many countries. Therefore, understanding the impacts of M&A on the efficiency and productivity changes in financial institutions has significant policy implications.

The rest of the chapter is arranged as follows: Section 2 makes a brief review of the literature about the effects of M&A on the efficiency and productivity changes in financial institutions. Section 3 describes the model and the variables used in the analysis. Section 4 and 5 gives the results of the estimation of efficiencies and productivities, respectively and make some analysis about them. Section 6 draws conclusions from the analysis.

5.2 A brief review of the literature about the effects of M&A on efficiency and productivity changes in financial institutions

Unfortunately researchers have no consensus about the effects of M&A on the efficiency and productivity changes in firms. Many researchers argued that M&A will improve the efficiency of the involved firms through: (1) technology transfers between the participant firms; (2) economy of scale and economy of scope improvement due to M&A; (3) reduction of overcapacity and redundant labors often carried out after M&A. On the other hand, M&A may increase the market power of the involved firms. This may increase the efficiency measured in value terms of the involved firms but have nothing to do with real productivity improvement. It may also reduce their motive for innovation. Thus it may decrease the growth rate of their efficiency measured in technical terms and also be disadvantageous to the interests of consumers.

There are by now numerous empirical researches in this field. For an early review of the literature about the effects of M&A on the efficiency and productivity changes of banking industry, see Berger, S.Demsetz et al. (1999). More recently literature includes: Hahn (2007) analyzed the effects of M&A on the efficiency of Austrian banks during the period of 1995-2002. Al-Sharkas et al. (2008) analyzed the impact of M&A on the efficiency of the US banks for the period from 1985 to 2002. Rezitis (2008) analyzed the effects of bank mergers on the efficiency and productivity changes for ten Greek banks from 1993 to 2004.

There are also several researches on the M&A among credit unions and other small financial institutions. For example, Grifell-Tatjé and Lovell (1996) analyzed the effects of M&A on the productivity changes of the Spanish saving banks from 1986-1991. Vennet (1996) analyzed the effects of M&A on the cost efficiency and other financial profit indicators of the credit union in Europe over the period of 1988-1993. Fried et al. (1999) analyzed the effects of mergers on the efficiency of credit union of the United States from 1989-1995. Garden et al. (1999) analyzed the effects of mergers on efficiency changes of the credit union in Australia during the period of 1992-1997. Haynes and Thomson (1999) analyzed the effects of M&A on the productivities of the building society in U.K. Devaney and Weber (2000) analyzed the effects of market structure on the productivity changes in the rural banking sector of the United States during the years from 1990-1993.

Corresponding to the inconsistence in the theoretical explanation, the conclusions of the empirical literature about the effects of M&A on the efficiency are mixed. Some have found no evidence (Grifell-Tatjé and Lovell 1996, Garden et al. 1999, Devaney and Weber 2000, etc.) that M&A had significant effects on the efficiency of the banks. Some (Fried et al 1999, etc.) found mixed evidences about the effects of M&A. some (Rezitis 2008, etc.) even reported negative effects of M&A on efficiencies. On the other hand, others (Vennet 1996, Haynes and Thomson 1999, Al-Sharkas et al 2008, etc.) have found positive relationship between M&A and efficiency.

There are also several papers which concern the small and middle financial institutions in Japan. These papers either directly analyzed the effects of M&A activities or used M&A as an important control variable. For example，Hoshino (1992) analyzed the effects of M&A occurred from 1969 to 1980 on some simple financial ratios of the small and medium financial institutions in Japan. He showed that, if judged by simple financial indicators, in general the small and medium financial institutions which had not experienced M&A behaved better than those which had. Only those which occurred between Shinkin banks and other kind of small and medium financial institutions have positive effects on the financial positions of the merged banks. He also found that the major reason behind the M&A was the desire for scale. Fukuyama (1996) analyzed the relationship between efficiency and scale of credit unions in FY 1992. He found the larger the scales of the credit unions were, the higher their pure and total efficiency were. Thus he supports M&A among large credit unions or credit unions with commercial banks. Harimaya (2004) analyzed the correlation between the efficiency and dividend policy of the Shinkin banks by using the data of FY 2002. He estimated the efficiency of the Shinkin banks measured in cost and profit using a stochastic frontier model. Then he checked the correlations between dividend policy and efficiency using factor analysis. He found no statistically significant relationship between efficiency and dividend policy. Horie (2010) analyzed the relationship between the operation areas and productivity changes in the

Japanese Shinkin banks during the period FY 2005-FY2008. In this research, I use the same selection of inputs and outputs as Horie when calculating the productivity changes of the Shinkin banks. I also follow his idea of weighted average when measuring some variables for the regional environments.

Measuring the effects of M&A on the efficiency or productivity changes is not an easy task. Some analysts (Fried et al. 1999，Rezitis, 2006, etc.) directly calculated the efficiency or productivity changes of the banks involved in M&A activities before and after the occurrence of M&A. However, this approach is problematic. M&A activities will cause discontinuity of data for the acquired or closed banks. As for the acquiring or the merging banks (some with a new name), the operating environment also has greatly changed; hence simple comparison of these banks before and after the merging is misleading. In our case of Japanese Shinkin banks some banks even experienced more than two M&A during the sample period, which makes the problem even more complex. Noticing this problem, some researchers (Ralston et al. 2001, etc.) compared the efficiency scores of the bank formed after M&A with the average efficiency scores of the merged and merging banks before M&A. Alternatively some researchers (Cooper, Seiford et al. 2006) advocate comparing a virtual bank which have aggregated inputs and outputs of the merged and merging banks to the bank formed after M&A. This is also not correct. For due to the nature of the measurement, the average radial efficiency score of two or more banks is not always equal to the efficiency score of the bank with aggregated inputs and outputs of these banks (Fried, Lovell et al. 2008).

In this chapter, I do not compare the efficiency and productivity changes of the banks which involved in M&A activities before and after the M&A incidences. Instead I compare the efficiency and productivity changes between the banks involved in M&A activities with those not involved before the sample period. All those which involved in M&A activities during the sample period are removed from the sample. The effects of M&A are inferred from the differences between these two groups.

I use a two-stage approach to analyze effects of M&A on the efficiency and productivity changes. To deal with measurement error and endogenous problems inherent in the second stage regressions, I use the algorithm suggested by Simar and Wilson (2007), but with a few alterations. To avoid the problem of bounding of the dependent variable, besides using a truncation model to estimate the effects of M&A on efficiency, I also use Malmquist index as the dependent variable. This avoids using the censored or truncation models and makes it possible only using OLS models, because Malmquist index is only low-bounded by 0 and its logarithm can even take negative value.

Finally, following Horie (2010), I use weighted regional economic data as the control variables. Compared to macroeconomic data, which are used in most of the other related papers, regional data are more suitable for the analysis of the regional financial institutions.

5.3 The methodology

I use a two-stage approach to analyze the effects of M&A on the efficiency and productivity change. It consists of two stages: in the first stage, an efficiency or productivity change measure is calculated. In the second stage, the estimated score is regressed on several environmental variables and a technique called semi-parametric bootstrapping was used to test the significance of the parameters of the model.

5.3.1 Approaches to encompass the environment variables in the analysis of the determination of productivities

Factors which contribute to the production process can be divided into two categories. One is those that directly take parts in the production; the other is those that can influence the productivity of the DMUs，but do not directly involve in the production. In the analysis of production process, the former is treated as inputs and is included in the production function. The latter is called the environment variables.

Analysts disagree on how the environment variables will influence the production process. Some assume DMUs under different environment conditions are facing different production functions (technologies); therefore they divide the total sample of DMUs into several different groups according to the environment conditions they each face, and then estimate a specific production function for each group.

The advantage of this approach is that a DMU is only compared with those DMUs which face similar operational environments; therefore the comparison is fairer. The problem of this approach is that this approach assumes the functional relationship between the productivity and environmental variables is not continuous. We can only compare the difference of productivities between different groups, but cannot find out the functional relationship between environmental variables and the productivity. This greatly reduced its policy attraction. Gilbert and Wilson (1998) , Garcia-cestona and Surroca (2008) have used this approach.

The second approach directly incorporates the environment variables into the estimation of productivities. In parametric approaches, this is realized by including the environment variables in the explanatory variables in the regression model. In non-parametric approaches such as DEA, it is realized by defining the environment variables as a special kind of inputs or outputs that cannot be controlled by the DMUs. Those environment variables which have positive effects on productivity are treated as inputs; while those environment variables which have negative effects on the productivity are treated as negative outputs.

This setting can analyze the functional relationship between each environmental variable and the productivity. It also avoids the problem of correlation between the input, output and environmental factors which I will explain later (Wang and Schmidt, 2002). But environmental variables hardly fit the concave requirement of inputs, which makes the analysis difficult. It is also difficult to determine the direction of influence of each environment variable in advance (Fried 2008). Analyses using this approach include Berger and Mester (2003), Al-Sharkas, et al. (2008).

The third approach believes all DMUs are facing the same production function. But the efficiency score a DMU got is influenced by the environmental variables. Based on this assumption, this approach first estimates a single production function for the whole sample. Next the estimated productivity scores are regressed on the environmental variables. This kind of analysis is also called as two-stage analysis. Papers using this approach include Fried et al. (1999), Garden (1999), Devaney and Weber (2000), Ralson (2001), Hahn (2007), Al-Sharkas et al. (2008), Casu and Molyneuxy (2003). This approach avoids the problems the above two approaches have. It is also more policy appealing since it directly shows the scale and significance of the impact each environmental variable has on the productivity. Thus, in my opinion it is better than the two other approaches.

There are also papers which utilized more than two approaches. For example, Horie (2010) utilized both approach 1 and 3. Girardone etc., (2004), Rezitis (2008) used both approach 2 and 3.

5.3.2 Problems with the second stage models

For a DMU *i*, $i = 1, 2, \cdots n,$ it can use $\mathbf{x}_i$ to produce $\mathbf{y}_i$, it has a efficiency measure of γ_i, γ_i is determined by a vector of environmental variables $\mathbf{z}_i$, then the basic two-stage model is:

$$\gamma_i = \mathbf{z}_i \boldsymbol{\beta}_i^{'} + u_i, \qquad i = 1, 2, \cdots n.$$

u_i is the error term.

Among those papers which used the two stage models, in the first stage most researchers used the efficiency score as the dependent variable. In the second stage, some analysts just used OLS (Devaney and Weber, 2000，Garden et al. 1999; Devaney and Weber, 2000; Ralson et al. 2001) to estimate the model. Some used GMM (Ataullah and Le 2006) or GLS(Isik and Hassan 2003) models.

However, for the second stage regression models using efficiency score as the dependent variable, several problems have to be solved. First of all, the dependent variable (the efficiency measurement γ_i) is bounded. For the input or hyperbolic-oriented Shephard distance, it is low bounded by 1. For the output-oriented Shepherd distance, it is low bounded by 0 and upper bounded by 1.

To solve the bounding problem of the efficiency scores, in the second stage most researchers used censored or Tobit (e.g. Hahn 2007) models to estimate the coefficients of the model. Simar and Wilson (2007) argued that truncation model may be more suitable in this case. Some analysts（Mester 1993，Girardone, et al. 2004, Fried et al. 1999, etc.）used logit models to avoid the problem. But logit model only suits for those efficiency measurements which is low bounded by 0 and upper bounded by 1. Another way to solve the problem is to use some ratio of the efficiency measure as the dependent variable. For like logarithm，ratios such as the Malmquist index are only randomly low bounded by 0; therefore models using these ratios as the dependent variable can be estimated by OLS.

Second, the dependent variable γ is not an observed variable. It itself is estimated from the sample in the first stage. By construction, the dependent variable $\hat{\gamma}$ is serially correlated. This is because γ_i is measured as the distance to the frontier and the entire sample *x* and *y* are involved in the construction of the frontier. Any DMU's change of *x* and *y* may cause change of the frontier, thus the efficiency scores of some DMUs. The fact that γ_i is correlated with any *x* and *y* also means u_i is also correlated with $\mathbf{z}_i$, because the choice of *x* an *y* of the DMU *i* is conditional on $\mathbf{z}_i$, thus here we also face a endogenous problem.

Furthermore, although there is significant theoretic progress in this field, the exact nature of serial correlation between γ 's and the correlation between u_i and $\mathbf{z}_i$ is still unknown. Recent researches (Kneip etal.2003) have found that the estimated $\hat{\gamma}$ is a consistent estimator of γ. Therefore asymptotically with $\hat{\gamma}$ converge to the real γ, both the serial correlation between γ 's and endogenous problem will disappear. However, $\hat{\gamma}$ converge to the real γ only at a rate of $n^{-\frac{2}{p+m+1}}$, that is:

$$\hat{\gamma}(x,y)-\gamma(x,y)=O_P(n^{-\frac{2}{p+m+1}})$$

Where *p* and *m* are number of inputs and outputs respectively. Therefore for the multiple outputs and inputs cases satisfying ($p+m>3$), the converging rate is slower than the normal converging rate $\sqrt{n}$ of traditional econometric models.

The third problem is that although $\hat{\gamma}$ is a consistent estimator, it is a biased estimator of real γ, Thus we also face a measure error problem in this model. This bias also disappears at a slower rate than that of the traditional econometric models.

The two and third problems are more difficult to deal with. Traditional econometric techniques cannot be used to handle these problems. Xue and Harker (1999) are the first to point out and deal with these problems using a semi-parametric bootstrap approach. Casu and Molyneux (2003) also attempt to solve the problems using a bootstrap approach. Simar and Wilson (2007) designed a more advanced semi-parametric approach for solving the problem. In this chapter, I use a semi-parametric approach similar to that suggested by Simar and Wilson to estimate the second stage model. Except choosing the efficiency score as the dependent variable, I also choose Malmquist index as the dependent variable. To further simplify the problem, I take log of the ratio. As mentioned above, ratios like the Malmquist index are only low bounded by 0. Their logarithms can even take negative values. Thus it avoids the bounding problem that efficiency measurement will face.

The Simar and Wilson approach has two algorithms. This chapter only uses the first one because the second algorithm involves using the estimated efficiency score (such as the Shephard distance) and its residuals to generate new input and output samples. It is much more complex. Although the second algorithm will correct the bias of the dependent variable, but it will also increase the variance of the model. It is also unsuitable in the case of the Malmquist index. For the Malmquist index, the algorithm can be described as follows:

Algorithm:

(1) Using the original sample production set $Q(x, y)$ to estimate the Malmquist

index $\hat{m}$.

(2) Using the traditional OLS method to estimate the model $\hat{m}_i = \mathbf{z_i}\boldsymbol{\beta} + u_i$ and get an estimation of the coefficient vector $\hat{\beta}$ and the random error term $\hat{u} \sim N(0,\sigma_u^2)$;

(3) Repeating the following step (3.1)-(3.4) *B* times to get a semi-parametric bootstrapping sample of $\hat{\beta}$:

(3.1) For each *i*, randomly draw u_i^* from $\hat{u}$;

(3.2) Calculate $\hat{m}_i^* = \mathbf{z_i}\hat{\boldsymbol{\beta}} + u_i^*$;

(3.3) Substitute $\hat{m}_i$ by $\hat{m}_i^*$ as the dependent variable, again using the traditional OLS method to estimate the model $\hat{m}_i^* = \mathbf{z_i}\boldsymbol{\beta} + u_i$ and get the estimation of parametric bootstrapping coefficient $\hat{\beta}^*$ as well as the bootstrapping random error $\hat{u}^*$.

(4) From the above bootstrapping procedure, we can get *B* samples of $\hat{\beta}^*$, then from these samples we can get a confidence interval for each element of the vector $\hat{\beta}^*$. The simplest way of calculating the confidence intervals of $\hat{\beta}^*$ is to use the quantiles of $\hat{\beta}^*$.

The required sample for semi-parametric bootstrapping is much less than the non-parametric bootstrapping used in the last chapter. Here I choose *B*=1000

For details about the problems in the second stage model and bootstrapping approach used in the model, see Simar and Wilson (2007)

5.4 The effects of M&A on efficiencies of the Shinkin banks

In this section, I estimate the effects of M&A incidences on the efficiencies of the Shinkin banks using FY 2008 as the sample year. I use the measurement of efficiency under the assumption of CRS as the dependent variable. As mentioned in

Chapter 3, efficiency measured under the assumption of CRS is a more composite measurement of efficiency. It includes efficiency measured under the assumption of VRS (pure efficiency) and scale economy.

5.4.1 Model specification

The regression equation is:

$$\ln crs = \mathbf{z}\boldsymbol{\beta}_c^{'} + \mathbf{u}_c \qquad (4.1)$$

Because hyperbolic-oriented efficiency measurement is low bounded by 1, all the dependent variables are truncated below by 0. Therefore the model is a truncation regression model.

I am especially interested in the effects of M&A incidences on the efficiency and productivity changes. I set dummies to capture the effects of M&A incidences. It may take times for the participant banks to integrate their operations and cultures. Therefore, the efficiency of the involved banks may at first decrease, and then gradually increase and finally the effects of the M&A incidences will disappear. To capture this dynamics I set 4 time dummies (date01, date02, date03 and date04): $Date_{it} = 1,\ t = 01, \cdots 04,$ if Shinkin bank *i* experienced M&A during fiscal year *t*.

To control for other factors that may influence the efficiency, I include several control variables in the second stage model. These variables are environment variables mentioned above. They can be divided into two categories: Internal factors and external factors.

Internal factors are bank specific variables that may influence the efficiency of the banks. In our model, these factors include:

(1) Total income (*asset*), as the indicator of bank scale. According to the firm theory，for each industry, there is an optimal scale of production. For firms smaller than this optimal level of scale, increasing the scale of production will increase their efficiency. On the other hand, for firms larger than this optimal level of scale,

increasing the scale of production will have negative effects on their efficiency. For regional institutions like Shinkin banks, their business is limited in a relatively small area; thus we should expect their optimum scale to be much smaller than national financial institutions. To capture this effects, I also add the square term of total income (asset08^2) into the model.

(2) Risk level indices. Since risk taking behaviors and profits of banks are correlated, these indices certainly will have effects on the efficiency of Shinkin banks. The paper chooses three indices as the indicators of risk taking.

(2.1) Risk adjusted capital adequacy ratio (CAR), the capital/risk weighted asset ratio defined by the Basle Accord. The sign of this coefficient is not easy to predict. Higher CAR ratio means a larger proportion of bank fund is not used in more profitable projects. This is detrimental to the efficiency. On the other hand, a strong capital position makes it possible for banks to take riskier projects. This will improve their efficiency. The results of empirical studies in this field are mixed.

(2.2) NPLs/ total loans ratio (*npl*); higher NPL ratio will reduce the efficiency of banks, thus this variable should have negative effects on the efficiency

(2.3) Loan loss provisions / total loan ratio (*lp*). The effects of this variable are similar to that of CAR.

(3) Indices of administration efficiency, an efficient administration will reduce the cost of production and strengthen the control of risk. Thus it will certainly have effects on the efficiency of Shinkin banks. Following Horie (2010), I include two indicators of administration efficiency in the model: one is the labor efficiency of the headquarter (LH). It is the ratio of the number of staffs in headquarters to the number of offices in the headquarters. The other is the organizational efficiency (NH). It is the ratio of the number of departments in the headquarters to the number of branches.

The two variables reflect the efficiency of headquarters. The higher the ratios, the higher the operating expense of the headquarters will be. However，they also reflect bank’s ability of supervision and risk control. Therefore the signs of these two variables are not predetermined.

External factors are those variables out of the control of the banks that affects the efficiency of the Shinkin banks. All of these factors are indicators of operating areas. Operating area is the economic and social environment in which a bank operates. As a kind of regional financial institution, Shinkin bank's activities are limited in a specific geographic area. Unlike large financial institutions which operate in a national scope, the economy of the region over which a Shinkin bank operates plays a key role in the determination of its efficiency (Horie 2010). In the model of this chapter, I use several economic and social variables as the indicators of operating area. These variables include:

(1) $Share_i$: the ratio of the number of branches of Shinkin bank *i* to the total number of branches of all financial institutions operating in the same region -an indicator of market power of Shinkin banks. As mentioned at Section 2, the role of market power in the determination of efficiency is not certain. On one hand, high market power may discourage a bank to improve its efficiency and technology, thus its efficiency may be low (market competition hypothesis); on the other hand, high market power also gives bank high price fixing ability; therefore its efficiency measured in value terms may be high (market power hypothesis). Due to these conflicting effects, the sign of market power indicators in the model is not easy to be pre-assumed.

(2) Indicators of regional economic activities. The model also includes three indicators for the level of regional economic activities:

(2.1) $income_i$: The taxable personal income of the region in which Shinkin bank *i* operates.

(2.2) $insti_i$: The number of enterprises of the region in which Shinkin bank *i* operates.

(2.3) $manu_i$: The value of manufactures of the region in which Shinkin bank *i* operates.

It is reasonably predicted that banks operate in areas which have high economic activities should have high level of efficiency; Thus these three variables all should have positive effects on the efficiency.

(3) Index of population characteristics. Population problem now is a key factor that influences the Japanese economy. Thus the model specially includes two indicators of population characters in the region:

(3.1) pop_i: The population density in the region in which Shinkin bank *i* operates. High density of population normally indicates more members and customers for the Shinkin banks. Therefore it should have positive effects on efficiency.

(3.2) old_i: The proportion of aged families to the total number of families in the region in which Shinkin bank *i* operates. High proportion of aged families is disadvantageous to the regional economy. Thus this variable should have negative effects on efficiency of the banks.

I use "city, town or village" (shi, mura or machi) as the unit of region. Since many Shinkin banks operate over more than one region, I use weighted average of regional statistics as the indicator of operating area. The weight is the ratio of the number of branches in city *i* of Shinkin bank *j* to the total number of branches of bank *j* in year *t*.

Since the scale of a region is different from each other, for quantity variables, it is questionable to treat each region equally. To avoid this problem, the three variables for economic activities is expressed in term of value per acreage.

5.4.2 The data

Data of the hyperbolic-oriented efficiency scores come from the estimation results from Chapter 4.

All of the data about the internal variables come from the Nikki database and the yearbooks "The Japan financial directory" published also by Japanese economic news. The descriptive statistics of internal variables are summarized in Table 5.1:

Table 5. 1: Descriptive statistics of the internal variables

	asset (billion)	lp (billion)	CAR (%)	npl (%)	LH (%)	NH (%)
Min.	38.41	-46.10	4.56	0.66	9.96	8.70
1st Qu.	121.50	-4.46	9.15	2.26	19.69	28.57
Median	250.50	-2.43	11.60	3.20	22.46	37.98
Mean	400.20	-3.94	12.68	3.49	22.85	41.97
3rd Qu.	486.10	-1.17	14.89	4.39	25.50	50.74
Max.	2619.00	-0.14	62.80	10.51	51.28	100.00
sd	443.18	5.22	5.81	1.75	5.16	18.34

The data about the market share are obtained from the yearbooks "The Japan financial directory." Other external data come from the "Regional Statistics Database" offered in the website of official statistics of Japan (www.e-stat.go.jp). Unfortunately for some external variables the data is not available for every year. Thus the end year for the variable *insti* is 2009. The end year for *pdensity* and *old* are 2010 respectively. Table 5.2 shows the descriptive statistics of the external variables.

Table 5. 2: Descriptive statistics of the external variables

	Share (%)	Income (billion)	Insti (billion)	Manu (billion)	Pop (million)	Old	crs
Min.	1.178	2.42	0.11	1.34	207.10	15.67	1
1st Qu.	17.1	8.95	0.39	12.77	712.70	22.05	1.104
Median	24.9	17.08	0.65	32.78	1281.00	24.27	1.163
Mean	26.49	50.56	1.74	55.80	2890.00	24.69	1.156
3rd Qu.	35.09	41.11	1.17	83.59	2632.00	27.14	1.201
Max.	66.26	509.60	38.77	376.40	16580.00	58.80	1.345
sd	13.502	84.87	3.43	58.88	3880.00	4.43	0.069

5.4.3 The results of the second stage model and their explanation

The results of the second stage model are summarized in table 5.3.

Table 5. 3: Estimation results of the second stage model for efficiency scores

	original	boot
(Intercept)	0.30287***	0.32242***
	(0.04314)	(0.04425)
asset08	-0.07201**	-0.07163**
	(0.03487)	(0.03745)
asset08^2	0.03566 ***	0.03650***
	(0.01295)	(0.01334)
lp	-0.00207*	-0.00265***
	(0.00111)	(0.00135)
CAR	-0.00395***	-0.00503***
	(0.00087)	(0.00109)
npl	-0.00377	-0.00451**
	(0.00250)	(0.00262)
LH	0.00210**	0.00227***
	(0.00075)	(0.00075)
NH	-0.00045*	-0.00051**
	(0.00027)	(0.00029)
date01	-0.03283*	-0.03895**
	(0.01808)	(0.01899)
date02	0.00387	0.00355
	(0.01159)	(0.01146)
date03	-0.01779	-0.01810*
	(0.01495)	(0.01498)
date04	0.01287	0.01368
	(0.02163)	0.02153
share	0.00060	0.00064*
	(0.00040)	0.00043
income	0.00005	0.00006
	(0.00016)	(0.00017)
insti	0.00620***	0.00627***
	(0.00188)	(0.00189)
manu	0.00004	0.00004
	(0.00009)	(0.00009)
pop	-0.01018**	-0.01114***
	(0.00391)	(0.00396)
old	-0.00417**	-0.00427***
	(0.00137)	(0.00136)

Note: a. ***: significant at 1% level.
**: significant at 5% level.
*: significant at 10 % level

b. figure in bracket is the standard error of the estimated coefficient.

We first check the four time dummies for M&A. From Table 5.3 we see that for original results only *date01* is negative and weakly significant (at10% significance degrees). But in the bootstrapping results, both *date01* and *date03* are negatively significant (at varied significance degree). This tells us that the M&A incidences occurred in FY2001 and FY 2003 have positive effect on the efficiency.

We then check other variables which may be related with the M&A activities. First of all, M&A activities will increase bank scale. The results show that coefficient of the variable "*asset*" (value of bank income) is negative and significant. Its square term is positive and highly significant. This indicates asset scale has positive effects on efficiency when the asset scale is below the optimal degree. But it will have negative effects on efficiency when the asset scale is above the optimal degree. This is consistent with the theory of scale economy.

M&A also may strengthen the market power of the acquiring or merging banks. The coefficient of the variable *share* (market power of the Shinkin banks) is positive. This means market concentration has negative effects on the efficiency. This is consistent with the hypothesis that banks with market power are reluctant in adopting new technologies and improving efficiency. However, it is only weakly significant for the bootstrapping results.

Though it is not our purpose, it is also interesting to check the results for other control variables. For internal factors, the sign of the variable "*lp*" (Loan loss provisions ratio) is significant in varied degrees for original and bootstrapping results. It sign is negative. The variable "CAR" (capital adequacy ratio) is also negatively and highly significant. This shows that high loan loss provisions ratio and CAR have positive effects on the efficiency. This is in line with the hypothesis that banks with adequate loan loss provision and higher CAR take riskier, but more profitable strategy. The variable "*npl*" (NPL ratio) is also negative, which shows *npl* has positive effects on the efficiency. But it is significant only in bootstrapping results. Though it may seem strange, this is in accord with the hypothesis that high profit banks are correlated with high risk.

We next check the effects of external variables on efficiency. The three indicators of economic activities are all positive. This means economic activities are negatively related with efficiency. This is not what we expected. It may be due to the fact that regions with high economic activity are regions with high competition. But only variable $insti_i$ (the number of enterprises) is highly significant in the equations.

For the two indicators of the population, the variable *pop* (the population density) and *old* (the proportion of aged families) are negative and significant at varied degrees for both the original and bootstrapping results. This means the population density and senility of the society have positive effects on the efficiency of the Shinkin banks. But the result for *old* is not what we assumed. However, some other researchers (Dietsch and Lozano-Vivas 2000) also got similar contradictory results. This may be because aged families normally have high saving, which is beneficial to the development of Shinkin banks.

5.5 The effects of the M&A on productivity changes

In this section, I further analyze the effects of M&A on productivity changes in Shinkin banks. The model used here are similar to the one used for the analysis of efficiency. Only now the dependent variable is Malmquist index and its components. Malmquist index based on hyperbolic measurement is only low-bounded by 0 and its logarithm can even take negative value. Therefore here we do not need to use the truncation model approach. The OLS approach is enough.

5.5.1 Model specification

Except the total scores, I also regress the components of the Malmquist index on the same environmental variables. Then we will get a system of regression:

$$
\begin{aligned}
&\ln \mathbf{m} = \mathbf{z}\boldsymbol{\beta}_m^{'} + \mathbf{u_m}, \\
&\ln \mathbf{e} = \mathbf{z}\boldsymbol{\beta}_e^{'} + \mathbf{u_e}, \\
&\ln \mathbf{t} = \mathbf{z}\boldsymbol{\beta}_t^{'} + \mathbf{u_t}, \\
&\ln \mathbf{s} = \mathbf{z}\boldsymbol{\beta}_s^{'} + \mathbf{u_s}, \\
&St. \ln \mathbf{m} = \ln \mathbf{e} + \ln \mathbf{s} + \ln \mathbf{t}
\end{aligned}
\tag{5.2}
$$

Utilizing the constraint, only three out of the four equations can be estimated. I omitted the equation for pure efficiency changes and estimate the remained three equation using OLS method. The coefficients of the omitted equation for pure efficiency and their significance levels can be deduced from the estimated models. The vector of coefficient of the equation for LnE is: $\boldsymbol{\beta}_e = \boldsymbol{\beta}_m - \boldsymbol{\beta}_t - \boldsymbol{\beta}_s$ and their corresponding variance is: $VAR(\boldsymbol{\beta}_e) = VAR(\boldsymbol{\beta}_m) + VAR(\boldsymbol{\beta}_t) + VAR(\boldsymbol{\beta}_s)$. With the estimate and its variances, it is easy to obtain its significance level. However, because the variances of the estimates are larger than those obtained by traditional method, its significant level will be underestimated. In this case bootstrapping method is a more accurate estimate. Through bootstrapping, we can obtain B estimate of each coefficient: $\boldsymbol{\beta}_e^* = \boldsymbol{\beta}_m^* - \boldsymbol{\beta}_t^* - \boldsymbol{\beta}_s^*$. The bootstrapping significant level can be estimated by using the quantile approach described above.

As for the explanatory variables, it is similar to those used in the model of efficiency described in section 4 of this chapter. Only now the log difference of the environment variables between FY 2005 and FY 2008 is used. I also add one variable LB (the efficiency of branches, which is average number of employees per branch) to the model. This variable reflects labor efficiency of branches. It should have negative effects on productivity changes.

5.5.2 Data in the second stage model

The dependent variables are the original estimated scores of Malmquist index and its components for Shinkin banks during the period FY 2005-FY 2008 obtained in Chapter 4. The descriptive statistics of internal variables are summarized in Table 5.4:

Table 5. 4: Descriptive statistics of the growth rate of the internal variables

Unit: %

	dlnLH	dlnLB	dlnNH	dlnAsset	dlnlp	dlnCAR	dlnnpl
Min.	-35.110	-40.260	-100.000	-9.457	-78.800	-59.750	-70.670
1st Qu.	-7.360	-7.536	-1.161	1.202	-30.800	-3.762	-36.970
Median	1.391	-1.925	4.555	3.832	-8.363	6.567	-24.310
Mean	5.899	-1.015	8.622	4.441	4.951	5.355	-19.360
3rd Qu.	13.110	4.235	17.040	7.801	21.050	16.730	-11.060
Max.	253.300	127.800	160.000	23.180	617.500	58.070	116.500
sd	25.783	12.719	29.218	5.408	67.440	18.043	27.219

From Table 5.4, we see that on average most of the variables have increased. The average asset of the banks increased about 4.44% during the period of FY 2005-FY 2008. This is consistent with the finding of the outputs showed in Table 4.1 in Chapter 4. For the cost part, the number of staffs per office of headquarter (LH) increased about 5.9%. The ratio of the number of departments in the headquarters (NH) on average increased about 8.6%. Risk adjusted capital adequacy ratio (CAR) on average increased about 5.36%. Loan loss provisions ratio (*lp*) on average increased by about 4.95%. Only two variables decreased. The number of employees per branch (LB) on average decreased about 1.02%. NPL ratio (*npl)* on average drastically decreased about 19.36%.

The explanatory variables are similar as those in the model for efficiency scores; except now they are logged and differentiated to suit the model here Table 5.5 shows the descriptive statistics of the external variables.

Table 5. 5: Descriptive statistics of the change of the external variables

Unit: %

	dlnshare	dlnincome	dlninsti	dlnmanu	dlnpop	dlnold
Min.	-49.110	-56.090	-56.690	-31.670	-51.950	-30.630
1st Qu.	-3.643	2.155	-0.212	3.803	-0.040	-0.020
Median	0.000	5.166	0.000	10.800	0.000	0.000
Mean	0.449	6.747	0.462	16.230	1.104	-0.125
3rd Qu.	3.583	8.953	0.000	18.800	0.000	0.000
Max.	102.500	388.300	202.500	656.500	311.600	15.580
sd	11.267	25.090	13.743	48.759	20.140	2.609

Table 5.5 shows that on average the market share of the Shinkin banks has only slightly increased by 0.449%. All three economic indicators for the regional economic activities have been increased on average. Among them the increase of manufactures are the highest. For the two indicators of population, surprisingly the average density of population has increased; while the average proportion of old people has slightly decreased by 0.125%. This is in contradiction with the trend of population growth in Japan. This means Shinkin banks have concentrated their business to area with more and younger population, which is favorable to their operations.

5.5.3 The results and their explanation

The results of the second stage model are summarized in Table 5.6.

Table 5. 6: Estimation results of the second stage model for Malmquist index

	LnM		LnT		LnS	
	Original	Boot	Original	boot	Original	boot
(Intercept)	0.03116	0.00377	0.03359	0.03361	0.00141	0.0013
	(0.00615) ***	(0.00569)***	(0.00204)***	(0.00197)***	(0.00373)	(0.00366)
dlnasset	-0.09079	-0.19115	0.06126	0.06126	-0.14101	-0.1383
	(0.05320)*	(0.05139)*	(0.02315)***	(0.02217)***	(0.03077)***	(0.03030)***
dlnlp	-0.0074	-0.01514	0.00564	0.00561	-0.00458	-0.00463
	(0.00786)	(0.00763)	(0.00336)*	(0.00331)*	(0.00454)	(0.00415)
dlnCAR	-0.03787	-0.0379			0.01432	0.01378
	(0.01550)**	(0.01520)***			(0.00895)	(0.00848)
dlnnpl	-0.00821	-0.00724	-0.00556	-0.00567	0.02416	0.02418
	(0.01208)	(0.01162)	(0.00533)	(0.00515)	(0.00699)***	(0.00661)***
dlnLH	0.00975	0.00983	-0.00759	-0.00778	0.01161	0.01144
	(0.01431)	-0.01374	(0.00632)	(0.00609)	(0.00822)	(0.00806)
dlnLB			-0.00947	-0.00948	-0.01342	-0.01312
			(0.01063)	(0.01007)	(0.01384)	(0.01353)
dlnNH	-0.01411	-0.01735	-0.00388	-0.00397		
	(0.00848)*	(0.00791)*	(0.00376)	(0.00367)		
date01	-0.01543	-0.01438	-0.00912	-0.00912	0.01039	0.01043
	(0.01332)	(0.01310)	(0.00589)	(0.00572)	(0.00770)	(0.00725)
date02	0.01067	0.01917	0.00192	0.00192	-0.00086	-0.00078
	(0.00897)	(0.00849)	(0.00394)	(0.00387)	(0.00518)	(0.00489)
date03	-0.00014	0.01379	-0.00811	-0.00784	-0.00089	-0.00077
	(0.01140)	(0.01097)	(0.00502)	(0.00472)*	(0.00659)	(0.00640)
date04	-0.01109	-0.0149	-0.01639	-0.01609	0.03002	0.02988
	(0.01684)	(0.01645)	(0.00739)**	(0.00730)**	(0.00974)***	(0.00949)***
dlnshare	0.03428	0.03697	-0.00939	-0.009	0.01941	0.01936
	(0.02327)	(0.02293)	(0.01019)	(0.00987)	(0.01353)	(0.01303)
dlny	-0.00966	-0.01931	-0.00415	-0.00415	-0.00776	-0.0077
	(0.01122)**	(0.01090)	(0.00486)	(0.00462)	(0.00657)	(0.00602)
dlninsti			0.00684	0.00697	-0.01264	-0.00631
			(0.00493)	(0.00466)	(0.00769)*	(0.00911)
dlnmanu	0.03326	0.03825			-0.0061	-0.0125
	(0.01586)	(0.01560)**			(0.00922)	(0.00754)*
Dlnpop	-0.00354	0.0229	-0.00979	-0.00982	0.01826	0.01846
	(0.01744)	(0.01663)	(0.00693)	(0.00677)	(0.01008)*	(0.00950)**
Dlnold	0.01484	-0.00826			-0.0078	-0.00793
	(0.02964)	(0.02703)			(0.02023)	(0.01933)
Adjusted R-squared	0.04283		0.05011		0.1326	

	LnE	
	Original	boot
(Intercept)	-0.00384	-0.00379
	(0.00748)	(0.00740)
dlnasset	-0.01103	-0.01389
	(0.06567)	(0.06661)
dlnlp	-0.00846	-0.00857
	(0.00968)	(0.00933)
dlnCAR	-0.05219	-0.05209
	(0.01789) ***	(0.01783) ***
dlnnpl	-0.02682	-0.02647
	(0.01494) *	（0.01462）*
dlnLH	0.00572	0.00708
	(0.01767)	(0.01796)
dlnLB	0.02289	0.0226
	(0.01745)	(0.01749)
dlnNH	-0.01023	-0.0108
	(0.00927)	(0.00896)
date01	-0.01669	-0.01652
	(0.01648)	(0.01614)
date02	0.00961	0.00943
	(0.01108)	(0.01067)
date03	0.00886	0.00894
	(0.01409)	(0.01381)
date04	-0.02472	-0.02389
	(0.02081)	(0.02034)
dlnshare	0.02426	0.02309
	(0.02878)	(0.02823)
dlny	0.00225	0.00211
	(0.01388)	(0.01336)
dlninsti	0.0058	-0.00066
	(0.00913)	(0.00989)
dlnmanu	0.03936	0.04581
	(0.01835) ***	（0.01692）***
dlnpdensity	-0.01201	-0.01209
	(0.02130)	(0.02079)
dlnold	0.02264	0.02354
	(0.03589)	(0.03481)
Adjusted R-squared		

Note: a．***: significant at 1% level.

**: significant at 5% level.

*: significant at 10 % level

b. figure in bracket is the standard error of the estimated coefficient.

All of the models have low adjusted R^2. The models of Fried et al. (1999), Ralson et al. (2001) and Horie (2010) also have this problem. Nevertheless some interesting conclusions can still be drawn.

We first check the four time dummies for M&A. From table 5.6, we see that they are all insignificant for the Malmquist index. However, when we look at the components, we see a different picture. For the technical efficiency，*date04* is significant (at 5% level). *date03* is weakly significant (at 10% level) only in the bootstrapping results. In the equation for scale efficiency, *date04* is highly significant (at 1% level). In the cases that the coefficients are significant, for LnT, the signs are negative, indicating M&A incidences in the period of FY 2001-FY 2004 have positive effects on the technical efficiency. This is in line with the market power hypothesis. However, for LnS, the signs are positive. This may indicate that many banks formed after M&A are oversized. The coefficient of *date04* is also much more significant in all equations. In the deduced results for LnE, the four time dummies are all insignificant. This tells that the M&A incidences have no significant effect on the changes of pure efficiency. The controversial effects of M&A incidences on LnT and LnS plus its insignificant effects on LnE may be the major reason why time dummies are all insignificant in the equation for LnM. As for the dynamics of the effects, the latest M&A cases have the most significant effects on these components.

We then check other variables which may be related with the M&A activities. First of all, M&A activities will increase bank scale. For the Malmquist index, the variable "*dlnasset*" (change of bank income) is only weakly significant both in the original and bootstrapping results (all at 10% level). However, it is highly significant in the equations for technical and scale efficiency (all at 1% level). It is insignificant in the deduced results for pure efficiency. In case of the total score and scale economy, the signs of the coefficients are negative. This indicates that the increasing of scale has positive effects on the changes of productivity and scale economy. On the other hand, in the equation for the technical efficiency, the sign of the coefficient is positive, which means bank scale has negative effects on technical efficiency changes. This supports the hypothesis that large banks lack the motive for technical progress. Combined with insignificance of the LnE, it is not surprising the variable is less significant in the equation for Malmquist index than in the equations for the two components. The results are consistent with the findings of Al-Sharkas et al. (2008) and Hahn (2007).

As mentioned in section 4, M&A may also cause the reduction of overcapacity and redundant labors. Thus they may bring out administration structure changes. For the three indices of administration efficiency, only dlnNH (number of department in headquarter/ number of branches) is weakly significant (at 10% level) in the equation for lnM, both in the original estimation and the bootstrapping results. Its sign is negative. This shows that a complex structure in headquarter is beneficial for the improvement of efficiency. Some variables are so insignificant in some equations that they are omitted from the corresponding equations.

The sign of the variable dLnshare (market power of the Shinkin banks) is positive in the equation for the total score, scale economy and the reduced results for the pure efficiency. This supports the hypothesis that market concentration has negative effects on the change of total productivity, pure efficiency and scale economy. On the other hand, in the equation for technical efficiency changes, its sign is negative, which is against the hypothesis the banks with market power are reluctant in adopting new technologies. However, they are all insignificant, whether according to the original estimation or the bootstrapping results. This means the change of market share has no significant effects on the productivity changes and its components.

As in the case of efficiency analysis, though it is not our purpose, we also check the results for other variables. For internal factors, the variable “dlnlp” (change of loan loss provisions ratio) is only weakly significant (at 10% level) in the equation for technical efficiency changes (dlnT). Its sign is positive. This means loan loss provisions have negative effects on the productivity changes. The variable “dlnCAR” (change of capital adequacy ratio) are significant for the Malmquist index and pure efficiency. In the equations which it is significant, its sign is negative, indicating that change of CAR ratio has positive effects on the productivity changes and pure efficiency. This supports the hypothesis that high capital position makes Shinkin banks take riskier but also more profitable strategies. The variable “dlnnpl” (change of NPL ratio) is significant only for the scale economy. Its sign is positive, indicating that increase of NPL ratio has negative effects on the change of scale economy. This is easy to understand, since banks with high *npl* ratio are constraint in their ability to increase their assets and exploit the economy of scale.

We next check the effects of external variables on productivity changes. For the three indicators of economic activities, the variable dlny (the change of taxable income) is only significant for the Malmquist index. In all equations, its signs are negative. This supports the hypothesis that Shinkin banks which located in an area with higher growth rate of personal income will have higher growth rate of productivity. The variable $d\ln insti_i$ (the change of the number of enterprises) is only significant for LnS (only for original results). Its sign in that equation is also negative. This supports the hypothesis that Shinkin banks which located in an area with higher growth rate of commercial institutions will have higher growth rate of scale economy. $d\ln manu_i$ (Change of value of manufactures) are significant in varied degrees for the equations of LnM, LnS and LnE. But in the equations for LnM and LNE, the signs of the variables are not what we expected.

For the two indicators about the population, the variable *dlnpop* (change of the population density) is significant only for LnS. However, in that equation its sign is positive. This means the growth of population density has negative effects on the growth of the scale economy. This is not what we assumed. The variable *dlnold* (the change of proportion of aged families) is insignificant for all the equations. In all equations its sign is negative. This means Shinkin banks in the area with high growth rate of aging population have high rate of productivity growth. Again this is not what we assumed. However, some other researchers (Dietsch and Lozano-Vivas, 2008) also got the same contradictory results for efficiency analysis.

5.6 Conclusions

In this section, based on the estimates of the hyperbolic-oriented efficiency and productivity changes of Shinkin banks in Japan during the period of FY 2005-FY 2008 obtained from the analysis in Chapter 4, we investigates the effects of M&A activities on the efficiency and productivity changes. It shows that merger incidents have no effects on the Malmquist index. But it has significant effects on the efficiency scores and two components (technical and scale efficiency) of the Malmquist index. It also shows that the merger incidents occurred in FY 2004 have stronger effects on these two components. This may be because it is closest to the examined period.

This chapter also finds that some other factors which are closely related to the M&A have significant effects on the efficiency and productivity changes. The variable bank scale, indicators of organizational efficiency and market share are significant at varied degrees in the model for the efficiency scores. More important, we find the variable of bank scales follow a quadratic form, indicating there is a optimal bank scale. In the model for productivity changes, the indicator of change of bank scale is significant for all the three equations and the indicator for change of organizational efficiency is significant for the equation of Malmquist scores.

Therefore from the results we may say that on the whole the M&A activities occurred during the early years of 2000s have significant effects on the efficiency and productivity change of the Shinkin banks. This may offer some support for the policies encouraging M&A activities among banks.

Chapter 6 Capital requirement and loan suppliers

—A theoretical background

6.1 Introduction

As discussed at the beginning of the dissertation, a major concern regarding banks is their ability to supply credit necessary for economic activities. To ensure that this ability is unimpaired, governments retain control over banks to mitigate risks. However, this regulation may contradict with the role of banks as loan suppliers.

One of the most important tools in controlling bank risk is the risk weighted capital adequacy ratio (CAR) regulation proposed by the Basel Committee on Banking Supervision. Since the publication of the first Basel Accord (Basel I) in 1988, CAR has been adopted in various countries. By 2006, the Accord was formally updated to a new version, Basel II, and a further version (Basel III), which was implemented in many countries in 2013.

There is an intensive debate about the effects of the CAR requirement on the risk taking behaviors of banks even before the birth of the Basel Accord. Various theoretical and empirical papers in this area have been published. In this chapter，I first make a short literature survey of related theories about the effects of capital control on bank loan supplies. Then, I develop a simple multi-sector model to explain the relationship among bank capital, loan supply, non-performing loans, and economic environment. The purpose of this chapter is to build a theoretical foundation for the empirical analysis in the next chapter.

6.2 A short literature survey of theories on the risk taking

behavior of banks under CAR regulation

This section provides an outline of related theoretical literature on the relationship between bank capital and loan growth.

6.2.1 B˚nk capital and loan supply

Different banks will have different reactions when the CAR requirement is imposed. The reaction depends on the capital conditions of each bank at the time of the introduction of the requirement. Obviously those banks that cannot meet the minimum requirement (under-capitalized banks) will have to take some action to satisfy the requirement.

Banks whose CAR requirement is binding can use four different methods or a combination of methods to solve the problem. These methods include: 1) enlarging their capitals by issuing new stocks or other debts recognized as tier 1 or 2 capitals; 2) decreasing the supply of new assets; 3) using credit arbitrage or short selling, such as securitization to remove accumulated credits; 4) re-arranging their assets structure to reduce the risk adjusted value of the total assets while the unadjusted value remains the same. The re-arrangement can be achieved by shifting their assets from those with higher risk weights to assets with lower risk weights. The first method will increase the numerator of CAR, whereas the remaining three methods will reduce the denominator of the ratio. Different methods will have different effects on the risk levels of under-capitalized banks. Different theories also have different assumptions on the actions of under-capitalized banks. These differences in assumption influence their results.

According to the conclusions of the models, theories on the effects of CAR regulation on loan supply (and other risky assets) of banks can in general be divided into two categories, namely, the “capital crunch” and the “risk shifting” schools (Saunders 2002).

The "capital crunch" hypothesis was first mentioned by Syron (1991). The "capital crunch" school predicts that the loan (and other risky assets) supply of a bank will decrease because of the tightening of CAR requirement. The reason behind this hypothesis is very simple. To fulfill the CAR requirement, banks with weak capital positions will be forced to either raise new capital in the capital market or reduce total risk assets. Since at most situations it is costly or even impossible for under-capitalized banks to issue new stocks, they will be forced to reduce the total volume of risky assets or shift their assets to those with lower weights (most likely government securities). If a large proportion of banks choose this option, a "credit crunch" may prevail. The phenomenon first caught the attention of economists in the early 1990s in the United States, stimulating a very hot discussion.

The "risk shifting" school argues that although under-capitalized banks can reduce their risk assets after implementation of the CAR requirement, they also may "shift" their assets to more risky ones because the risk weight rules of the Basel Accord are not perfect. Assets with different levels of risk may be assigned the same risk weight, leaving sufficient room for the operation of "risky shifting." Therefore, the net effects of introducing CAR regulation are ambiguous. In some cases, risk shifting effects may overwhelm the risk reducing effects, thereby increasing the total level of risks of banks.

Economists have used different methods to analyze the risk shifting effects. They can be roughly grouped into three categories, including the complete market approach, portfolio selection approach, and game theory approach.

Rochet (1992) employs the complete market model to analyze the effects of CAR requirement on the risk taking of banks. In a complete market setting, CAR regulation has no influence on the risk taking attitude of banks, regardless whether any kind of explicit or implicit (e.g., in the form of government rescue of default banks) deposit insurance is present. In the absence of deposit insurance, market discipline will force banks to take a cautious attitude toward risk. In this case, CAR regulation is unnecessary. On the other hand, if deposit insurance exists, depositors will not pay attention to the risk level a bank takes. In this case, all bank investments will flow to a specific risky asset with the highest risk among the same expected returns (Rochet, 1992, proposition 1). If CAR requirement is imposed, then the level of risk may be reduced. Nevertheless, the banks will still choose the assets with the highest risk. In other words, no risk shifting will occur (Rochet, 1992, proposition 3).

Therefore, the complete market approach is not very useful in analyzing the risk shifting behavior of banks. Besides, as Repullo (1992) appointed out, any complete market assumption will immediately raise the question why the existence of banks is necessary.

Some analysts treat banks as a portfolio manager and use the portfolio selection theories (mainly the mean-variances analysis) to study the effects of CAR regulation on risk taking behavior of banks. The models are developed from the portfolio models of Pyle (1971), Hart and Jaffe (1974), and was first used by Kahane (1977). Kohen and Santemero (1980), Kim and Santemero (1988), and Rochet (1992) further developed the model. Among them, Rochet synthesized the other models, and in my opinion, his model is the best. Below, I will mainly use the model developed by Rochet to discuss the major conclusions of the portfolio selection school.

Assume that:

a) Banks behave like a portfolio manager;

b) Equity capital C cannot be increased ($\Delta C = 0$) in the model, implying that banks cannot increase their CAR by adding new capitals. This is because it is hard to model the behavior of the prices of the bank stocks[9].

c) No credit arbitrage or short-selling is allowed; otherwise CAR cannot be binding.

d) The value of capital at the end of period ($\tilde{K}_1$) is normally distributed with mean μ and variance σ^2

We can then apply the mean-variances analysis to solve the problem of portfolio selection. Rochet first analyzed the case with no bankruptcy (unlimited liabilities). In

[9] In the models of Kohen and Santemero (1980), Kim and Santemero (1988), banks are allowed to issue new stocks at a fixed exogenous price. The volume of C will be decided endogenously along with other variables in their models. If C can be increased, unless its cost is higher than the riskless assets, the CAR condition cannot be binding.

this setting, if there is no CAR regulation, the efficient set of asset choices is a straight line called market line. In this case, Rochet proves that the default probability of a bank is a decrease in the function of its capital adequacy ratio (Proposition 6).

When CAR requirement is added, the efficient set of asset choices is a combination of a portion of the "market line" and a part of the non-decreasing upper bound of set $\frac{\tilde{K}_1}{<\alpha, x>} > \bar{CAR}$ (α is the risk weight, *x* is the vector of values of assets, and $\bar{CAR}$ is the required minimum level of CAR). The latter is less steep than the market line (Rochet, proposition 8). Thus, the effects of CAR requirement are the kinking at some point of the market line toward the riskless assets.

If the risk weight α is proportional to the β in CAPM theory, then the hyperbole is reduced to a horizontal line of $\mu = \bar{\mu}$. In this case, the capital constraint banks will have a corner solution at the intersection of the two lines, implying a less risky portfolio. Those who choose a portfolio alone the market line are not constrained by the CAR requirement. Their risk taking decisions will not be affected.

However, If α is not chosen according to β, for those banks which are capital constrained, the total value of risk weighted assets will be reduced if additional capitals cannot be placed; on the other hand, these banks will "shift" their assets to more risky assets to maintain their profit rate at the level before the imposition of the CAR requirement. Thus, a "risk shifting" phenomenon will occur. The net result will be ambiguous. In some cases, the banks become riskier. Kim and Santomero (1988) obtain the same results.

If we further add an assumption that the liabilities of banks are limited, The value function that banks need to maximize become：

$$W(\mu, \sigma) = U(\mu, \sigma) - CN(-\frac{\mu}{\sigma}),$$

where $C \geq 0$ is the fixed cost of bankruptcy, $N(-\frac{\mu}{\sigma})$ is the normally distributed possibility of bankruptcy.

The new objective function W is now neither uniformly convex nor concave with respect to σ. Rochet (1993) demonstrates some severely under-capitalized banks may tend to be risk-loving for reasons similar to the complete market setting. Thus, a moral hazard problem may appear. Whether a bank will prefer risk-loving will depend on the time substitution of its preferences. In fact, for $-\frac{u''(x)}{u'(x)} \leq \alpha$, if $K < 1/\alpha$, W is increasing in σ (Rochet 1992, proposition 11.) Similar results are obtained by Kahane (1977), Kohen and Santemero (1980)

Portfolio selection approach produces some interesting results; however, this approach has a major weakness. CAR as suggested by Basel Accord I is based on credit risk, which deals mainly with the possibility of default, and not market risk, which comes from the fluctuation of bank assets. Repullo (1992) points out that portfolio selection approach may better suit the analysis of non-bank financial institutions. However, with the implementation of Basel II and Basel III, this approach may become suitable for analyzing the risk taking behavior of banks.

Some economists use game theories, particularly the "moral hazard" concept to analyze the risk taking behavior of banks and attempt to predict the consequences of CAR restriction imposed on banks.

To understand the "moral hazard" problem, we should first understand the importance of franchise value in the risk taking behavior of banks (Caprio and Summers, 1993). Franchise value can be treated as the capitalization of profit earning ability. Bank franchise value is a special case because most countries have no free entry into the banking industry. Setting up a bank requires special government authorization.

Since the franchise value of a bank is the capitalization of its monopoly profit earning ability. It is positively related to its profit rate. The franchise value of a bank is decreases as its profit rate decreases. The franchise value disappears when the bank ceases to operate. Thus, when the franchise value of the bank is low, for the same reasoning of low capital value, the bank will follow a more risky strategy, since in this situation, the bank has less to lose and much more to gain. In the extreme case, when the value of capital $K \le 0$, the bank will become extremely risk-loving. This condition is a typical case of “moral hazard” phenomenon.

Franchise value can shed some new light on the relationship between risk taking behavior of banks and their CARs. If the capital of a bank increases faster than its assets, on the one hand it will let the bank be more prudent, since it now has more its own money to lose if the investment fails; on the other hand, it will lower down the profit rate per share, thus the bank will be tempted to take more risks.

Hellman et al. (2000) use an infinite repeated game model to analyze the risk taking behaviors of banks. In each period, at first stage, a bank can freely choose its level of capital *K*, but at a cost higher than the returns of its prudent investment. This assumption guarantees that the CAR requirement can be binding in some cases. The bank then offers an interest rate to attract deposits in a competitive market. After raising the deposit, at the second stage the bank allocates its fund into two projects: one prudent project with a constant return α and another gamble project that yields an expected return $\theta\gamma + (1-\theta)\beta < \alpha$ (θ is the possibility that the gamble yields a return $\gamma > \alpha$, and thus the possibility that the gamble yields a return $\beta < \alpha$ is $1-\theta$). If the bank chooses the prudent project, the game will continue; if the bank chooses the gamble project and the gamble is successful, then the bank obtains higher return $\gamma > \alpha$ and continues the game. If the gamble fails, the bank will close and its franchise value will disappear (no forbearance of failed banks in the model).

From social efficiency perspective, banks should choose the prudent project because $\theta\gamma + (1-\theta)\beta < \alpha$. However, banks may have to pay an inefficient high deposit interest rate because of the severe competition in the deposit market. If no capital is required, Hellman et al. show that the only sub-game equilibrium is that banks pay a high deposit interest rate and invest in the risky project. Adding the CAR requirement can solve the problem; however, CAR would have to be set at relatively high level because of the controversial effects of capital as mentioned above.

Deposit interest rate ceiling is another policy choice. However, as the regulation Q of the United States in the early 1970s has shown, the ceiling may become binding and cause a disintermediation problems or force banks to find ways inefficiently to shirk the regulation. Hellman et al. suggests a policy that combines the two regulations. In this way, the ceiling can be set at a high level such that it will not be binding easily and the CAR requirement can be set at a much lower level than the case with only CAR regulation. Hellman et al. proved that there always exists a combination of the two policies that dominate the policy using only CAR regulations in the sense of Pareto optimum (proposition 4).

In the model of Hellmans, et al, there is no "bail out" by the government if the risky project fails, but in real world, this bail out is an important factor in the risk taking decisions of banks, as seen in Japan before 1995 and in former socialist countries. Berglof and Roland (1995) develop a "gamble for bail out" model to describe the risk taking behavior under the "soft budget" environment. In this model, the government has an interest to continue the failed investment. When an investment fails, the government would rescue it. The rescue would only be in the form of recapitalization of the banks. Thus, if banks believe that the government would rescue a loan project that fails, they would provide loans to the project although they are aware ex ante the project would fail. Berglof et al. argue that recapitalization under soft budget constraint cannot let under-capitalized banks to be prudent in their investment decisions; therefore, this scheme could not solve the lack of capital problems. Only a combination of recapitalization and hard budget constraint would succeed.

Agénor and Pereira da Silva (2012) develop a model to analyze the role of bank capital in an imperfect market. Under the assumption of imperfect market, bank capitals could be used as a signal of the safety of banks to the depositors. Well capitalized banks could attract deposits with lower interest rates than poorly capitalized banks. Thus, bank capital could also influence the bank loan supplies indirectly through loan cost in addition to its direct influence on loan supply. This effect exists even for banks whose capital requirement is not binding. Thus, they find that bank capital is pro-cyclical regardless whether the capital is constrained or not.

6.2.2 Dynamics of the reaction of banks to the tightening of CAR requirement

Some economists have attempted to analyze the dynamics of the reaction of banks to the tightening of CAR requirement. For example, Shrieves and Dahl (1992) develop a simultaneous equation model (SEM) to analyze the relationship between bank capital and the adoption of risk-based capital standards. The model consists of two equations:

$$\Delta CAP_{i,t} = \alpha(CAP^{*}_{i,t} - CAP_{i,t-1}) + E_{i,t}$$
$$\Delta RISK_{i,t} = \beta(RISK^{*}_{i,t} - RISK_{i,t-1}) + S_{i,t}.$$

In the model, $CAP_{i,t}$ and $RISK_{i,t}$ are the actual (observed) capital position and risk level for bank *i* in time *t*, respectively. $CAP^{*}_{i,t}$ and $RISK^{*}_{i,t}$ are the targeted capital position and risk level, respectively. $E_{i,t}$ and $S_{i,t}$ are the external shock on capital position and risk level, respectively. To cope with unexpected investment opportunities, banks normally set the target capital level above the minimum CAR set by supervisors.

Calem and Rob (1999) analyze and quantify the dynamic portfolio choice of banks by using a empirical database from the US banking industry over the period from 1984–1993. They point out that banks with different capital positions react differently to capital-based regulation. They find a U-shaped relationship between capital position and risk-taking; under-capitalized banks took maximum risk. Banks take less risk as its capital rises. However, when its capital continues to rise, the bank would undertake more risks again. Severely under capitalized banks assume higher risks because the costs of bankruptcy have shifted to deposit insurance companies. Well-capitalized banks assume higher risks because of their higher ability to absorb losses. In addition, the paper also examined the comparative effects of a flat versus risk-based capital regulation.

Hyun and Rhee (2011) set a dynamic multi-sector model to analyze the choice of banks when faced with capital constraints. Except for the high cost of issuing new shares, Hyun and Rhee (2011) offer another explanation as to why banks would rather reduce loans instead of issuing new capitals. They prove that the incumbent shareholders of banks prefer to reduce loans rather than issue new capital because issuing new capital would dilute the earnings of incumbent shareholders.

Meh and Moran (2010) create a dynamic stochastic general equilibrium (DSGE) model to analyze the role of bank capital in economic activities. The model includes a sector of final goods and a sector of intermediary goods and householders. Bankers and investors are involved in the production of capital goods. The role of bank capital is to solve the moral hazard problems of banks. The existence of capital enables depositors to believe that banks will use the funds they obtain properly. Thus, a sudden decrease in bank capital will affect the ability of the banker to attract funds from investors because of reduced credibility, causing a significant decline in loan supply. This decline will result in procyclical bank capital.

6.2.3 Comparison of the effects of Basel I and Basel II

Since its publication in 1988 (Basel I), various problems in Basel I have surfaced during its implementation, and thus a revised version (Basel II) was proposed in 2004. From 2006 to the end of 2011, Basel II was implemented in G20 countries. Since its publication, Basel II has attracted the attention of many analysts. The focus is on whether the new Accord is more procyclical than the old one. Different from Basel I, which based risk weight on a simple classification of debtors (firms, governments of developed countries, governments of developing countries, private lending, mortgage lending, etc.), Basel II uses a much more complex and flexible method to determine the risk weight assigned to an asset. Risk shifting has more flexibility because the credit classification is much more differentiated than the old accord. Furthermore, its calculation of risk weighting depends mainly on the credit rating by external credit rating institutions or internal judgment of banks. Because these ratings usually have a procyclical tendency (an asset will obtain a higher rating in boom and a lower rating in bang), it is highly likely that under Basel II bank loan supply is more pro-cyclical than under Basel I.

Jacques (2008) establishes a simple one period model to analyze the effects of Basel II on the loan supply. In the model, bank assets are divided into three categories, namely, loans with low risk (L^L), loans with high risk (L^H), and a security with further lower risk (*S*). Different kinds of assets have different required ratio of capitals (γ^L, γ^H, and γ^S). They prove that under the condition of Basel II, capital constrained banks will reduce more high risk loans than under Basel I ($\gamma^L = \gamma^H = \gamma^{88}$, where γ^{88} is CAR under Basel I). Meanwhile, the Banks will reduce low risk loans less or may actually increase them. Thus, Basel II will reduce the total level of bank risks, but its effects on total loan supply are ambiguous. Its result will depend on the risk weights assigned to different assets, the degree of competition in the bank market, and the capital buffers (surplus capital above the minimum requirement).

Agénor and Pereira da Silva (2012) also analyze the differences in the effects between Basel I and Basel II. In their model, risk weights under Basel II are correlated with the risk premiums that banks charge upon borrowers. Under Basel II, loans to high risk borrowers will be assigned a higher risk weight compared with those to low risk customers. However, the demand for the former will also be lower because of high risk premium. The latter will counteract the procyclical effects of bank capitals. They concluded that under Basel II, bank capital is less procyclical than under Basel I.

The 2007 sub-prime debt crisis in the United States and the European sovereign debt crisis re-ignited interest among analysts on the relationship among Basel II, loan supply, and economic cycles. For example, Heilpern et al. (2009) analyze the role of Basel II in the “credit crunch” that occurred after the 2007 sub-prime debt crisis in the US. As mentioned previously, one of the techniques to avoid capital requirement constraint is securitization. Securitization can reduce the value of risk assets and save capital for more loans while allowing banks to earn considerable profits from sales. The risks are shifted to financial markets. This change in operation strategy encouraged the increase in mortgage rates in the US and pushed the price of houses into an unsustainably high level.

Bank assets under the Basel II framework are valued at market prices and their risk weights are determined by the possibility of default (PD). Both are high in the period of asset bubbles, but may decline drastically when the bubbles burst. Therefore, when the prices of houses drastically went down in 2007, banks suffered huge losses. This decline weakened the capital positions of banks, resulting in the decrease in bank loan supply.

After the 2007 sub-prime debt crisis, G20 agrees to begin the implementation of Basel III at the beginning of 2013, with the implementation targeted to be completed by the end of 2019. Compared with Basel II, Basel III classifies some banks as "globally systematic important banks." These banks are required to add 1% to 2.5% more capital in addition to the standard 8% requirement. The required common stock CAR is also raised from 2% to 4.5%. First-tier CAR is raised from 4% to 6%. This strengthening of capital requirement is feared worldwide because it could lead to capital shortage of banks and a "credit crunch" at a time when the world economy is in recession and loans are urgently needed. US Federal Reserve estimated that the largest 19 bank holding companies in US need to raise US$ 50 billion to meet the new requirement. The US Federal Reserve announced in November 2012 the indefinite postponement of the implementation of the new accord because of fear of the "credit crunch."

6.3 Bank capital, loan supply, and economic growth

In this section, we define a simple multi-sector model consisting of a series of simultaneous functions to explain the interactions between bank activities and economic growth. The model is a discrete time dynamic model.

6.3.1 Agents in the model and their behavior function

This model has three agents. Government is excluded to simplify the model.

6.3.1.1 Firms

Firms use a technology represented by a simple AK model to produce domestic products. That is:

$Q_{t+1} = AK_t$,

where Q_{t+1} is product in year *t+1*, K_t is the stock of the capital in year *t*, and A is the constant ratio of products to capital.

To simplify the model, we first assume that no other financial market exist other than the banking market. The profits of firms are also distributed to shareholders. Bank loan is the only source that firms can use to finance their investments. Then,

$$
\begin{aligned}
K_t &= L_t \\
\Delta K_t &= I_t - \delta K_t = \Delta L_t - \delta L_t \\
&= NL_t - (\chi + \delta)L_t - DBL_t
\end{aligned}
$$

where L_t is the bank loan in year *t*, δ is the depreciation rate of capital, χ is the rate of repayment of loans, NL_t is the new loan supplied by banks in time *t*, and DBL_t is NPL deleted in time *t*.

We assume at time *t+1*, a proportion of β_t of L_t turned to be bad. We further assume that β_t is determined by the equation:

$$\beta_t = a + \frac{b}{2}(L_t - \bar{L})^2,$$

This means β_t is convex in L_t. It reaches the lowest point of $\beta_t = a$ at point $\bar{L}$.

The production function then becomes

$$Q_{t+1} = A(1-\beta_t)L_t = A\{1-[a+\frac{b}{2}(L_t - \bar{L})^2]\}L_t.$$

The basic reasoning under the equation is that a proportion of bank loans will be non-performing regardless the prudence of banks, which is captured by the constant *a*. However, another proportion of the total NPL is correlated with *L* in quadratic form. When *L* is below the critical level $\bar{L}$, the increase in *L* will reduce the ratio of bad loan (β). However, after $\bar{L}$, the increase in *L* will raise β.

In this case, Q_{t+1} is no longer linear in *L* (*K*), but cubic in *L* (*K*). Thus, the model is no longer endogenous. Although from the point of a single successful investment project, the production function is linear to loan (capital), from the total loan (capital), it is cubic in loan (capital). Therefore, the production function is first convex in loan (capital), then after some critical value of loan (capital), it turns to be concave. This assumption is in accord with the normal assumption on production function.

6.3.1.2 Banks

Banks follow the basic balance of payment equation:

$D_t + C_t = L_t + R_t$,

where D_t is the volume of deposit, C_t is the bank capital in time *t*, and R_t is deposit reserve of the banks in time *t*.

With the absence of government and financial market, loan (*L*) is the only asset other than the deposit reserve that a bank can hold. To simplify the model, we assume that the bank market is a complete competition market. Then, there is an infinite demand for any bank and interest rate is exogenously determined. Loan supply has two constraints. In normal conditions, loan supply will be constrained by the volume of deposit D_t and the minimum deposit reserve rate (d); however, banks are also required by the supervisor to meet the minimum CAR requirement (k). Thus, the loan supply constraint becomes $L_t \leq \min\{(1-d)D_t, kC_t\}$.

Similar to other kinds of firms, banks maximize their profit Π_t, which, in our model now, is simply the interest gap between loan and bank debit:

$$\Pi_t = r_L L_t + r_D(C_t - D_t),$$

where Π_t is the net revenue of banks in period *t*, r_L is the interest rate of loan, r_D is the interest rate of deposit. For simplicity, we assume the rate of return of deposit reserve R as zero and bank capital can obtain the same interest rate as the bank deposit. Initially, we assume that banks can only increase their capital by profit.

5.3.1.3 Family

Families receive all the net products from firms in the form of wages and stock dividends. They also receive deposit interest income and dividends from banks. After counter-deleting the deposit interest received by the family and paid by banks as well as the loan interest paid by firms and received by banks, the family income in time *t* (Y_t) is:

$$Y_t = Q_t - \delta K_t + r_D C_t - \Delta C_t - DBL_t.$$

Because we assume banks have no other sources to raise capital, the only source banks can increase their capital and delete NPLs is to bank profit. Thus ΔC_t and DBL_t should be deduced from the bank dividends received by families.

Families save a fixed percentage of their income. Given that no other financial market exists and we also assume that no government exists in the model, the only financial instrument a consumer can purchase is bank deposit, Then, $S_t = D_t$ and $\dot{D}_t = \dot{S} = sY_t$, where S_t is the accumulated savings at time *t* and D_t is the volume of deposits the bank system attracts in time *t*.

6.3.2 Optimal behavior of banks

The maximization problem of banks can be expressed as follows:

$$\underset{L_t}{Max}: \Pi_t = r_L(1-\beta_t)L_t + r_D(C_t - D_t)$$

$$\begin{aligned} St:\ & \beta_t = 1 - b/2(L_t - \overline{L})^2 \\ & L_t \leq (1-d)D_t, \\ & C_t \leq \kappa L_t \end{aligned}$$

Case 1: All inequality constraints are unbinding.

We first check the case that all inequality constraints are unbinding. The first order condition will be:

$$\begin{aligned} & \frac{d\Pi_t}{dL_t} = r^L(1 - a - \frac{b}{2}(3L_t^2 - 2\overline{L}L_t + \overline{L}^2) = 0 \\ & \Rightarrow (1 - a - b(3L_t^2 - 2\overline{L}L_t + \overline{L}^2) = 0 \\ & \Rightarrow 3L_t^2 - 2\overline{L}L_t + \overline{L}^2 - \frac{1-a}{b} = 0 \end{aligned}$$

The second-order derivative of the object function is:

$$\frac{d^2\Pi_t}{dL^2} = -r^L\frac{b}{2}(6L_t - 2\overline{L}).$$

It is easy to see that Π_t is convex when $L_t < \frac{1}{3}\overline{L}$ and becomes concave after it.

Thus it has a maximum when $L_t > \frac{1}{3}\overline{L}$.

Solving the first-order equation, we obtain:

$$\begin{aligned} L_t^* &= \frac{2\overline{L} \pm \sqrt{4\overline{L} + 12(\overline{L}^2 - \frac{1-a}{b})}}{6} \\ &= \frac{\overline{L} \pm \sqrt{4\overline{L}^2 - 3 \times \frac{1-a}{b}}}{3}. \end{aligned}$$

Thus, the maximum level of loan at time *t* is:

$$L_t^* = \frac{\bar{L} + \sqrt{4\bar{L}^2 - 3 \times \frac{1-a}{b}}}{3}.$$

At this maximum level of loan, the deposit reserve will be:

$$R_t^* = D_t - L_t^*.$$

The deposit reserve rate is:

$$d_t^* = R_t^* / D_t = 1 - L_t^* / D_t > d.$$

Given that the minimum deposit reserve requirement is not binding, the bank system will have surplus deposit. Banks will adjust their loan supplies through control of the surplus reserves. In this situation, as long as $d < d_t^*$, the minimum deposit reserve rate policy, one of the most important monetary policy instruments, will be ineffective.

CAR at this maximum level of loan is

$$k_t^* = C_t / L_t^* > k.$$

There are also surplus capitals in the bank system. The minimum CAR requirement policy will also be ineffective.

Note L_t^* is also the optimal level of loan we can obtain from the solution of optimal problem for firms. Therefore, a balance of loan supply and demand exists at the optimal level.

Case 2: Minimum deposit reserve rate requirement is binding:

If the minimum deposit reserve rate requirement is binding, then $d > 1 - L_t^* / D_t$, then the maximum loan level (L_t^d), deposit reserve (R_t^d), and the corresponding level of CAR (k_t^d) are as follows:

$$L_t^d = (1-d)D_t < L_t^*$$
$$R_t^d = dD_t$$
$$k_t^d = C_t / L_t^d = C_t / (1-d)D_t \ .$$

In this situation, the loan is determined by the level of deposit, which is below the optimal level of loans. If the government reduces the minimum deposit reserve rate, loan supply will increase. The efficiency of banks as well as that of firms will also increase. Monetary policy (minimum deposit reserve rate requirement) will be effective. However, the minimum CAR requirement policy is still ineffective because the CAR minimum requirement is not binding.

Case 3: Minimum capital requirement is binding

If the minimum capital requirement is binding, that is, $k > C_t / L_t^*$, then the maximum loan level L_t^k, deposit reserve (R_t^k), and corresponding level of reserve rate (d_t^k) are expressed as

$$L_t^k = C_t / k < L_t^*$$
$$R_t^k = D_t - L_t^k = D_t - C_t / k$$
$$d_t^k = R_t^k / D_t = 1 - \frac{C_t}{Dk} > d \ .$$

The deposit requirement is not binding. Therefore, the monetary policy will be ineffective, but the CAR requirement policy is effective. Relaxing the CAR requirement will increase the loan supply. Furthermore, it will also improve the efficiency of banks and reduce NPL ratio.

6.3.3 Dynamics of the system

Suppose that at the beginning banks have an initial deposit D_0 and capital C_0. If in these initial levels of deposit and capital, $(1-d)D_0 > L_t^*$, $k / C_0 > L_t^*$, that is, neither the minimum deposit reserve rate nor CAR requirement is binding, then bank loan supply will remain at the optimal level $L^* = L_t^*$ during the entire period. Given that loan supplies are constant during the entire period, $\dot{L}_t = 0$, then

$$\begin{aligned} &\dot{L}_t = NL_t - \chi L_t - DBL_t = 0 \Rightarrow \\ &NL_t = \chi L_t + DBL_t \end{aligned}.$$

This means the new loans supplied by banks at any time *t* should equal the repaid loans plus deleted NPL loan at the same time. Since the repaid loans χL_t can be self financed from former loans, the only problem is whether sufficient profit is available for the purpose of loan deletion, that is, $\Pi_t \geq DBL_t$. If this condition is satisfied, then banks can maintain the maximum level of loans. If we do not consider the discount factor, NPL should be deleted as soon as it appears, that is, $DBL_t = \beta_t L_t$ ($\beta_t L_t$ is the total volume of NPLs). If profits are inadequate to delete NPLs, then the capital stock has to be used to delete NPLs. If the capital position is below the minimum required level after deletion of NPLs, then we turn to case 3.

Now let us consider case 2 in which the minimum deposit reserve requirement is binding. If only minimum deposit reserve requirement is binding at the beginning, then $L_0 = (1-d)D_0$. Given that saving rate *s* is fixed in our simple model, loan supply will increase as *t* increases as long as *s* > 0. From forward induction it is easy to see:

$$\begin{aligned} L_t &= (1-d)D_t = (1-d)(D_0 + \sum_{t=1}^{T_d} sY_t) \\ &= (1-d)(D_0 + \sum_{t=1}^{T_d} sA(1-\beta)L_{t-1} - \delta L_{t-1} + r_D C_t - DBL_t). \end{aligned}$$

The new loan supply is limited by increased deposit that can be used for loans. Thus, the new loan in time *t* will be

$$NL_t = (1-d)\Delta D_t = (1-d)sY_t$$

Similar to case 1, we assume that $\Pi_t \geq DBL_t = \beta_t L_t$. Then, the dynamics of loan in time *t* will be

$$\Delta L_t = (1-d)sY_t - \chi L_t - \beta_t L_t.$$

If the capital is below the minimum required level after deletion of NPLs, then we also turn to case 3.

As the deposit D_t increase, at a critical point T_d, $L^d_{T_d} = L^*$, the minimum deposit requirement will become unbinding. After this critical point, the loan supply will be maintained at the optimal level. Correspondingly, the savings will remain at a constant level $\Delta S = sA(1-\beta)L_{t-1} - \delta L_{t-1} + r_D C_t - DBL_t$

Next, we consider case 3. In this case, the minimum CAR requirement is binding at the beginning. If banks cannot obtain new capital from the market, then they can only increase their capital through profit. A tradeoff between loan supply and NPL deletion occurs in this situation. Deleting NPL reduces the profits that can be used to enhance bank capital. When bank capital requirement is binding, this will reduce bank loans. As the bank loan is already below the optimal level L^*, reducing bank loans will diminish bank profit. Thus, there would be lesser profits left to increase loans or to delete NPLs, resulting in a vicious cycle.

Therefore, in this situation, the deletion of NPL should be postponed. If there is no maximum NPL ratio requirement, then at the beginning no NPL should be deleted. The loan supply is expressed as

$$L_t = C_t / k = \sum_{t=1}^{T_L} (C_{t-1} + \Pi_t) / k$$
$$= \sum_{t=0}^{t} [C_{t-1} + r^L (1-\beta_t) L_t + r^d (C_t - D_t)] / k$$

Similar to case 2, as long as $\Pi_t > 0$, C_t will increase as t increases. As C_t increases to the critical point T_c, $L_{T_k} = L^*$, the minimum CAR requirement becomes unbinding. After this critical point, the loan supply will remain at the optimal level, and surplus profit should be used to cancel accumulated NPLs. Therefore, the dynamics of DBL_t after critical point T_c is

$$DBL_t = \Pi_t .$$

If, however, there is an upper constraint of NPL ratio (*npl*) set by the supervisor, then NPLs will accumulate to a level after some time T_{npl} that this constraint is binding. In this case bank profits should be used to delete NPLs to satisfy this upper constraint. Then the dynamics of the DBL_t will be:

$$DBL_t = (1-npl)NPL_t .$$

The equation for capital and its dynamics is expressed as

$$\Delta C_t = \Pi_t - (1-npl)NPL_t$$
$$C_t = C_{t-1} + \Pi_t - (1-npl)NPL_t .$$

Loan supply is

$$L_t = C_t / k = \sum_{t=1}^{T_L} (C_{t-1} + \Pi_t - (1-npl)NPL_t) / k$$
$$= \sum_{t=0}^{t} [C_{t-1} + r^L (1-\beta_t) L_t + r^d (C_t - D_t) - (1-npl)NPL_t)] / k$$

It is also possible that both the minimum deposit reserve and capital constraint are binding. Assume that the deposit reserve constraint is stricter than that for CAR, that is, $L_0^d = (1-d)D_0 < L_k^0 = C_0 / k$. Then at first the deposit reserve constraint will be effective. As deposits increase with time, the deposit reserve constraint will become loose at the critical point of time T_d. After the critical point, the minimum CAR requirement constraint will become effective. C_t also increases with time, at the critical point T_c, it also will become unbinding. After this critical point, the loan supply will remain at the optimal level. It is also possible that the CAR constraint is stricter than the deposit reserve constraint ($T_c < T_d$), then the CAR constraint will be effective at first.

In the case that the minimum CAR requirement is binding, if we give up the assumption that banks cannot obtain capitals from the financial market; then banks can avoid the minimum CAR constraints by issuing new stocks in the market. It is easy to see that the optimal capital level C^* corresponding to the optimal loan level L^* is:

$$C^* = kL^* .$$

If banks initially have binding capital C_0, then the capital that banks need to issue (C_t^s) is determined by

$$C_t^s = kL^* - C_0 .$$

However, in this case we also need to include a security market in our model. We maintain the assumption that the security market does not offer direct financing for investment. In this case, the savings in year *t* will be divided between bank deposits and the bank stocks.

$$S_t = D_t + C_t^s \quad \text{and} \quad \dot{S} = sY_t = \dot{D}_t + \dot{C}_t^s .$$

If the deposit at the beginning is not binding, then we are facing the case with no constraint is binding. Therefore, loan supply will remain at the maximum level L^* during the entire period. If, however, the deposit becomes binding after the required bank capital is fully deducted from the savings, the surplus savings that can be used to purchase bank capital is:

$$C_t^s = sY_t - L^* / (1-d) .$$

The dynamics of capital becomes:

$$\dot{C}_t = \Pi_t - DBL_t + C_t^s = \Pi_t - DBL_t + S_t - L^* / (1-d) \Rightarrow$$

$$C_t = C_0 + \sum_{t=1}^{T_c} [\Pi_t - DBL_t + sY_t - L^* / (1-d)]$$

Until at the critical point T_c, the minimum CAR requirement becomes loose and the situation turns to case 1.

Chapter 7 Capital requirement and loan suppliers of regional banks in China

In this chapter, based on the theoretical analysis in Chapter 6, an empirical study on the relationship between capital requirement and loan supplies is presented by using a sample of city commercial banks in China during the period of 2005–2008. In early 1998, the People's Bank of China (PBC, the central bank of China) began to cancel quantity control on bank loans. Henceforth until 2006, this policy was gradually substituted with CAR management. China also began to implement Basel III in 2013. This chapter tries to check whether this policy change has significant influence on loan growth of regional banks in China. Or put it another way: does the capital condition become a significant constraint on loan growth of regional banks in China?

The structure of this chapter is organized as follows. Section 1 provides a brief literature survey of empirical studies on the relationship between bank capital and loan supplies. Section 2 describes the specifications of the model and panel data used in the model. Section 3 provides the results and explanations of the regression. Section 4 provides some policy recommendations and concludes the chapter.

7.1 Empirical analysis on the relationship between capital position and loan supplies

This section outlines empirical literature on the relationship between capital positions and loan supplies. These researches try to testify the various hypotheses offered by theoretical models mentioned in the previous chapter.

7.1.1 Effects of bank capital on loan supply

As mentioned in Chapter 6, there are mainly two different theories about the relationship between bank capital and loan growth: “the capital crunch” school and the “risk shifting” school. Many empirical studies tries to testify which school is correct. This task is not easy because business cycle factors (or demand factors) may influence the loan market simultaneously. During periods of recession, firms tend to be pessimistic on the economic future, and their desire for loans will decrease. Banks may also become more risk-averse and reluctant to supply loans. The opposite may also be true during periods of economic boom. However, disentangling the influences of these factors and isolating the effects of capital regulation on loan supply can be a difficult task.

One way to overcome this shortcoming is to utilize panel data and examine the relationship between capital and loan supply under different economic and capital regulation environments. If, under similar economic circumstance, the relationship between capital and loan supply becomes tighter during the period when capital regulation is tougher, then we may say that capital regulation has a significant influences upon bank supplies.

Some analysts follow an indirect method to explore the effects of bank capital on loan supply by analyzing the demand factors (economic fundamentals) in the loan market. The reason behind this approach is that if the demand factors cannot explain the slowdown of bank loans fully, then the supply factors must have something to do with it. Mosser and Steindel (1994) use this approach to analyze the “credit crunch” of the United States during the period of 1989–1992. They develop demand models for four forms of credits and find that all four models significantly over-predict the real values of loans ($Y < Y^*$, where Y and Y^* are actual and predicted levels of loan, respectively) during the “credit crunch” period, implying that the growth rate of credit in this period is exceptionally low. Hence, some supply factors must be responsible for that result.

Agenor et al. (2004) use a similar indirect approach to test the credit crunch hypothesis for the East Asia financial crisis in 1998. They develop a demand model for excess liquid reserves of banks, and then use the model to see whether there are "involuntary" accumulation of liquid reserves (reserves that are much larger than what the demand model will predict) , If this is the case, then we would say there is a credit crunch. By using this method, they find that Thailand was indeed undergoing a credit crunch.

The demand side approach has a major restriction. The approach implicitly assumes that no structural change occurs in the loan demand model. Otherwise, to use historical data to predict the loan demand is inappropriate. If the effects of the demand cannot be predicted correctly, then the effects of supply cannot also be estimated correctly. For example, some analysts point out that new inventory management technology has resulted in a considerable decrease in the required volume of inventories in many industries and is partly responsible for the decline in the demand for liquidity credit of firms in the United States. The approach is also unsuitable for analyzing loan growth in a transforming economy because these economies usually have enterprise reforms operating simultaneously with bank restructuring.

Other economists directly analyze the factors that influence loan supply. For example, Bernanke and Lown (1991) use a reduced form model for loan supply to estimate the effects of CAR regulation on loan supply during the "credit crunch" period in the US in 1989–1992. They assume that the influence of changes in capital regulation can be inferred from the coefficient of capital in a regression of bank loan growth on bank capital and various control variables for loan demand. They find the effect of falling bank capital on lending to be small but statistically significant; suggesting that in most regions, capital shortage only has a modest effect on the availability of loans. They also examine the other types of credit extension, which have declined since the beginning of the recession, and show that falling CAR is also a major factor in the slowdown of these assets.

Lown and Wenninger (1992) use a cross-sectional regression model to analyze the role of the banking system in the United States in 1989–1992. The regressions prove the link between bank CAR and loan growth is stronger in during the period of 1989–1991 than in 1988 (the coefficient, its t-value, and R^2 ratio are much higher), particularly when only CAR is included in the model.

Horiuchi and Shimizu (1998) use a supply model to analyze the relationship between loan supply and CAR for Japanese banks before 1995. They find that the growth rate of loan is negatively related to the capital/asset ratio of Japanese banks. Woo (2003) uses a method similar to that of Lown and Wenninger (1992) to analyze the effects of bank restructuring on loan growth during the Japanese financial crisis. Woo uses a panel data of Japanese banks and runs a series of cross-sectional regression (loan growth regressed against CAR) for each year from 1991 to 1997. He finds a negative relationship (significant) between bank loan growth and CAR from 1991 to 1994, indicating that Japanese banks did not pay attention to their capital positions; however, the coefficient of CAR after 1995 became positive and R^2 ratio increased significantly, indicating that banks became increasingly aware of their capital positions.

Otchere et al. (2003) conducts a case study on the privatization of the Common Wealth Bank of Australia (CWBA). They find that although privatization has significantly increased the capital adequacy ratio of CWBA, privatization also stimulated the banks to follow a more aggressive strategy, and thus their NPLs increased.

Watanabe (2010) develops a bank loan supply model to analyze the choice of customers of Japanese banks during the "credit crunch" period from FY1995 to FY2001. He found that when they face a capital constraint, banks will cut loans to those customers which is relative safe, but will continue to offer loans to those which have difficulty in paying the outstanding loans to prevent these loan gone bad. This will increase the risk of the banks assets.

All the above empirical analyses are carried out using the example of one country. Chiuri, Ferri, et al. (2002) collect an international panel sample from 15 emerging economies with heterogeneous bank and economic conditions, and use a model similar to that of Peek and Rosengren (1995) to testify the "capital crunch" hypothesis. They find that bank credit growth is negatively related to the tightening of capital requirement, particularly at less well-capitalized banks and bank-based emerging countries and the negative impact had been larger for countries that had enforced CARs in environments of currency/financial crises. Even in countries with relative sound bank systems, the result still held.

To isolate the effects of bank capital on loan growth and economic growth from other demand and supply factors mentioned above, some analysts have explored the effects of foreign bank conditions on the domestic loan market because foreign bank loans are exogenous in the determination of domestic loan supply. Peek and Rosengren (2000) conduct an interesting empirical study on the effects of the Japanese bank crisis on the US real estate market. They use a panel data model that is differentiated by the characteristics of markets. Because the loan supply of Japanese branch is external for the U.S market, this case study offers a good opportunity to isolate the supply and demand factors that affect loan growth. Furthermore, because Japanese bank activities are concentrated in a few markets, they are able to calculate the effects of loan shrinkage in Japanese branches on US real estate growth. The research finds that Japanese banks significantly reduced their loan activities in the US after 1995 because of their problems in Japan. The withdrawal of Japanese banks has significant effects on the growth of construction sector in the US because these Japanese banks are deeply involved in the US real estate loan market.

Similarly, Brana and Lahet (2009) analyze the effects of capital requirement of Japanese banks on the Asian crisis in 1997. They find that heavily under-capitalized Japanese banks (due to the bursting of the bubble) in the early 1990s had to shrink loans from Asian countries. This move had significant effects on the outbreak of the Asian crisis in 1997.

Paravisini (2008) analyzes the effects of a government funding program for regional banks in Argentina between January 1995 and December 2001 to determine the effects of bank capital on loan supplies. He uses the funding a bank receives as proxy for the change of capital. The endogenous problem between bank capital changes and loan growth is avoided because funding is exogenous. He finds that regional banks in Argentina are capital constrained, and increasing their capital has positive effects on their loan supplies.

7.1.2 Dynamic interaction among bank capital, loan growth, and economic growth

One of the big problems in the analysis is the endogenous relationship among bank capital, loan growth, and economic growth. The rise of CAR will strengthen the ability of banks to supply loans. This will raise the economic growth rate. High economic growth will reduce NPL ratio. Loan growth will also increase the demand for capital. If banks can raise capital in the market, then the bank capital will also increase. However, according to Hellmann, Murdock et al. (2000), after some optimal levels, the speed of NPL growth will exceed loan growth. This will increase NPL ratio and reduce the positive effects of loan growth on economic growth. The increase in NPL will also absorb bank capital and reduce CAR. This endogenous relationship between capital requirement and economic growth may form a vicious cycle and result in the so-called "credit crunch" phenomenon during the period of economic recessions.

Several papers have attempted to capture this endogenous relationship. Some researchers use simultaneous equation models (SEM) first suggested by Shrieves and Dahl (1992) to analyze the endogenous relationship between the change in bank capital and risk. For example, Jacques and Nigro (1997) use a sample of US commercial banks to analyze the effect of CAR on the change in capital and risks from the end of 1990 to the end of 1991 (the first year of CAR implementation in the US). They find that the implementation of CAR significantly increased the capital adequacy ratio and decreased the risk levels of sample banks. But they found no significant difference in this effect between capital adequate and inadequate banks. Rime (2001) conducts a similar analysis for a sample of Swiss banks from 1989 to 1995 and find that Swiss banks with capital positions close to the minimum CAR attempt to raise their CAR to the minimum level. They also found that CAR regulation has a positive and significant influence on capital-asset ratio, but not on risk weighted CAR. This result implies that "risk shift" is not a serious problem in the

Swiss banking system. They believed that this is because in Switzerland raising capital through stock market was not very difficult for banks. Roy (2004) addresses the effect of the enforcement of the 1988 Basel Accord on banks in G10 countries. The paper analyzes the adjustments in capital and portfolio risk in banks from G10 countries over the period from 1988–1995. The result suggests that banks with capital positions close to the Basel standards generally increased their capital adequacy ratio without any offsetting increases in portfolio risk (except in France and Italy). The evidence also proves the U-shape relationship hypothesis as proposed by Calem and Rob (1999).

Other analysts use VAR models to analyze the endogenous relationship among bank capital, bank assets, and economic growth. For example, Hancock, Laing et al. (1995) establish a VAR model using the quarterly data of a sample of US banks to analyze the endogenous relationship among bank capital, bank assets, and economic growth. They find that banks normally take a few quarters, sometimes even more than a year, to adjust their assets following a shock to their capitals. Because of the sluggish of the adjustment, the "credit crunch" caused by capital shocks lasted longer than the adjustment of capitals. Banks with inadequate capital reduce more credits than banks with adequate capital. They also find that small banks adjust their assets slower than large banks.

7.1.3 Comparison of the capital requirement between Basel I and Basel II

Ever since the announcement of Basel II, in order to analyze the effects of the new Accord on the bank loan supply, some analysts tries to compare the effects of capital requirements between Basel I and Basel II.

Antão and Lacerda (2011) compare the capital requirements for credit of non-financial firms under Basel I and Basel II using a sample of Portuguese banks. They find that under any reasonable assumption on the coefficients used in the calculation function of risk weights defined by Basel II, in general the capital requirements for the credit of non-financial firms are lower under Basel II than Basel I, especially for the credits of large corporations and small and medium firms are classified as retail credits.

Using a sample of ten Norwegian banks, Andersen (2011) simulates the effects of Basel II on the capital positions for a wider scope of risk assets than that of Antão and Lacerda (2011). The simulation is based on a system of simultaneous equations. Andersen finds that the risk weights of bank assets increased in the scenario of recession. This increase has negative effects on bank asset growth. Consequently, the increase has also has negative effects on economic growth. He also finds that bank capitals decreased during the recession period because of heavy loan losses. This further aggravates the problems and forms a vicious cycle.

7.1.4 Empirical literature on the relationship between bank capital and loan supply in China

Several studies investigate the relationship between bank capital and loan supply in China. Zhao and Wang (2007) use a model similar to that of Peek and Rosengren (1995) and Woo (2003) to analyze the effects of capital position on loan supply for a sample of 12 Chinese banks during the period from 1995–2003. They find that capital position has no significant effects on the loan supplies of sample banks.

Dai, Jin et al. (2008) analyze the effects of capital management on bank loan and the monetary policy in China during the period of 1998–2005 by using data on the four largest national banks (“The Big Four”). The most innovative point of the study is that they estimate a simultaneous model that includes both supply and demand equations for bank loans. They find that the effects of the monetary policy are asymmetric because of capital supervision in China. Capital supervision is more effective in dealing with overheating problems than with regression, because the capital requirement makes banks more reluctant to supply loans.

Wu and Zhou (2006) consider the endogenous relationship between capital and risk of commercial banks in China by using a simultaneous model similar to that of Shrieves and Dahl (1992). Their database is a sample of 14 large and medium banks in China between 1998 and 2004. They find that the implementation of CAR regulation in China has negative significant effects on the risk of banks, whereas its effects on the change in capital are positive but insignificant.

Wang and Wu (2012) use an unbalanced panel data of Chinese commercial banks in the period of 1998–2009 to analyze the effects of capital supervision on bank loans. They establish a reduced form model to analyze loan growth in China. The model includes a dummy variable for the implementation of capital supervision, dummy variables for banks with capital constraints (8% < CAR < 10%) and those with inadequate capital (CAR < 8%), variables for level of capital (first tier CAR, CAR, capital buffer), and several control variables. Their results show that holding other factors constant, bank loan growth rate declines by 9% because of the implementation of capital requirement. The capital level is positively related to loan growth rate over the entire sample period. On average, the loan growth rate of capital constrained banks is less than that of the capital adequate banks, whereas the growth rate of capital inadequate banks is less than that of capital constrained banks (insignificant, however).

Xu and Chen (2009) develop a dynamic stochastic general equilibrium (DSGE) model and use quarterly macroeconomic data from 1993 to 2005 to simulate the dynamic relationship between bank credit and economic fluctuation in China. They find that credit shocks can explain most of the fluctuations of short term consumption, loans, and real money balance. However, their model has no capital constraint. Only deposit constraint is included.

In this chapter, I also develop an SEM that considers the influences in both directions. This approach offers a clearer description of the complex interrelationship among bank conditions, loan supply, and economic environment. The model in this chapter differs from the models mentioned above in that it considers directly the endogenous relationship among NPL, loan growth, and capital changes. As mentioned above, loan growth will influence NPL growth, and the accumulation of NPL will negatively affect the capital position of the bank. The weakening of bank capital position will in return affect bank loan growth.

I also make great effort to enlarge the sample. It is well known that the sample of banks with detailed data available is very small in China, especially before 2008.Instead of only relying on almanacs or available databases such as Bankscope, I use the annual reports of Chinese banks collected from the website. Compared with other related papers on the banking system in China, to my knowledge, the sample in this research is the largest. Different from most studies in the field, which use time series data or bank panel data combined with macroeconomic data, I use a bank panel database combined with regional economic data. I believe regional economic data are more suitable in the analysis of regional financial institutions. Furthermore, to measure the economic environment each regional bank faced more precisely, I use the weighted average of regional economic data for banks that operate in more than one region.

7.2 Model specification and description of data

7.2.1 Specifications of the model and its variables

SEM is used to analyze the relationship between bank loan behavior and bank capital conditions. This approach considers the count-effects of loan growth on bank conditions. Thus, the analysis will produce more accurate estimates of the coefficients of the model.

The model is defined as:

$$
\begin{aligned}
d\ln loan_{it} &= b_{01} + b_{11}\, car_{it} + b_{21} npl_{it} + b_{31} d\ln deposit_{it} + b_{41} d\ln gdp_{it} + b_{51} gov_{it} + b_{61} foreign_{it} + \mu_{it1} \\
npl_{it} &= b_{02} + b_{12} car_{it} + b_{22} d\ln loan_{it} + b_{32} npl_{it-1} + b_{42} d\ln gdp_{it} + b_{42} ni_{it} + b_{52} gov_{it} + b_{62} foreign_{it} + \mu_{it2} \\
car_{it} &= b_{03} + b_{13} npl + b_{23} d\ln loan + b_{33} car_{it-1} + b_{43} d\ln asset_{it} + b_{53} d\ln capital_{it} + b_{63} d\ln gdp_{it} + b_{73} gov_{it} \\
&\quad + b_{83} foreign_{it} + \mu_{it3}
\end{aligned}
\tag{7.1}
$$

where

$d\ln loan_{it} = \ln loan_{it} - \ln loan_{it-1}$, $Loan_{it}$ is the total nominal value of outstanding loan of bank *i* in year *t*;

$d\ln capital_{it} = \ln capital_{it} - \ln capital_{it-1}$, $capital_{it}$ is the owner's equity, which is the sum of common stock, reserve, and after tax incomes in year *t*, corresponding to the concept of tier 1 capital in the Basel accord;

$d\ln deposit_{it} = \ln deposit_{it} - \ln deposit_{it-1}$, $deposit_{it}$ is the total volume of deposits of bank *i* in year *t*;

$d\ln asset_{it} = \ln asset_{it} - \ln asset_{it-1}$, $asset_{it}$ is the total volume of assets of bank *i* in year *t*;

car_{it} is CAR of bank *i* in year *t*;

npl_{it} is the ratio of NPL to the total loans of bank *i* in year *t*;

ni_{it} is the ratio of net income to total assets of bank *i* in year *t; and*

$d\ln gdp_{it} = \ln gdp_{it} - \ln gdp_{it-1}$, gdp_{it} is the weighted GDP in year *t* in regions where bank *i* operates.

Two variables that reflect the ownership structure of the banks are also included. gov_{it} is the proportion of stocks of bank *i* in year *t* owned by the government. $foreign_{it}$ is the proportion of stocks owned by foreign investors of bank *i* in year *t*. To reflect the dynamics of the system, one year lagged *car* and *npl* (car_{it-1} and npl_{it-1}) are also included as explanatory variables.

In the model, dln*loan*, *car* and *npl* are defined as endogenous variables. The interest of this study is the coefficient of CAR (*car*) in the first equation. Thus, the study attempts to testify whether the capital position has significant effects on loan growth. The null hypothesis is that *car* has no effects on loan growth. Hence, the following hypothesis is checked: H_0: $b_{11}=0$.

If this hypothesis is rejected, according to the theory of "capital crunch," in the equation for dln*loan*, banks with higher *car* should have higher loan growth rate. Therefore, its coefficient should be positive. Hence, the hypothesis: H_1: $b_{11}>0$ is also tested. By contrast, the "risk shift" school will predict that banks with higher *car* should have lower loan growth rate. Thus, its coefficient is assumed to be negative. The corresponding hypothesis is: H_3: $b_{11}<0$.

For the same reason, banks with higher *npl* ratios will have low loan growth rates according to the "credit crunch" school, but will have high loan growth rates according to the "risk shift" school. Thus, their coefficients should also be negative and positive, respectively.

In the equation for *npl*, according to Stiglitz and Weiss (1971), if we assume that bank loans have already reached the optimal level, higher loan growth will cause higher *npl* because the numerator will increase faster than the denominator. By contrast, if bank loans are under the optimal level, loan growth will reduce *npl*, because the numerator will increase slower than the denominator. The time lag is also an important consideration. If NPL has long time lag, NPLs will not be added into the stock of NPL in the period it originated, but in later periods. Then, high loan growth in time *t* may decrease *npl* in the current period, but will increase it in later periods. Due to these considerations, the coefficient of *dlnloan* cannot be predetermined. As for *car*, according to the "capital crunch" school, high car enables banks to assume a riskier operation strategy and afford higher *npl*. If so, its coefficient is presumed to be positive. However, the "risk shift" school maintains that lower car will let banks more willing to gamble, hence, cause more *npl*. Therefore, the sign of its coefficient should be negative.

In the equation for *car*, higher dl*nloan* will increase the demand for capital, thus its coefficient should be positive. Banks use capital to reduce *npl*, thus *npl* is assumed to be negatively related to *car*. The sign of the coefficient should be negative.

Each equation also has several exogenous variables. As the most important source of loan, the growth rate of deposit (*dlndeposit*) is included in the equation for dln*loan*. According to the classical deposit multiplier theory, this variable is assumed to be positively related to dln*loan*. For Chinese banks, deposit has a special meaning. According to the "Commercial Bank Law" published in 1995, the ratio of loan to deposits of a commercial bank cannot exceed 75%. I also include AR(-1) terms of *car* and *npl* in the equations for *npl* and *car,* respectively. Both AR(-1) terms are assumed to be positively correlated with their current terms; that is, banks with high *npl* (*car)* in a year will also have high *npl* (*car)* in the next year. Net income ratio (*ni*) is included in the equation for *npl* and *car*. The profit of the bank can be used to delete NPL. Therefore, *ni* is assumed to be negatively related to *npl*. By definition, *car* should be positively related to *ni*. Capital growth rate (dln*capital*) and asset growth rate (dlnasset) are included in the equation for *car*. Same as *ni*, dln*capital* should be positively related to *car* by definition. As with loan growth, high asset growth will increase the demand for capital. Thus, dln*asset* should be positively related to *car*. The two ownership indicators appear in all three equations. Banks with a higher share of foreign ownership tend to be more conservative in loan supplies. Therefore, *foreign* should be negatively related to *dlnloan* and *npl*, but positively related to *car*. By contrast, governments have a tendency to favor more loan supplies and care less about the safety of the loan (particularly local governments in China). Thus, *gov* should be positively related to *dlnloan* and *npl*, but negatively related to *car*.

All of the above variables are internal variables reflecting the characteristics of a bank. *dlngdp* is an external variable used as a proxy for the economic environment a bank faces. Economic environment can influence the demand for bank loans by customers, as well as the willingness of banks to supply loans. A regional financial institution usually operates in a limited market called "operation area." Thus, macroeconomic data are not suitable proxies for the economic environment of a regional bank. If a regional bank operates in a single region (a county, a province, etc.), then we can use regional economic data as proxy for the economic environment. In fact, several analysts use dummy variables for regions as proxy for economic environment (Hancock et al., 1995). However, in recent years, many regional banks in China have expanded their business over the region from which they originated. For these regional banks, using single region data is no longer suitable. Thus, I use the weighted regional GDP as proxy for the economic environment of a regional bank. As in Chapter 5, I use the share of the number of branches in a region over the total number of branches a bank owned as the weight for the bank.

Among the three endogenous variables in the model, dln*loan* is determined by the banks, whereas *npl* and *car* are partly determined by the economic environment and government. Thus, the system is well specified. At a quick check, each endogenous variable has at least one unique instrumental variable. Therefore, all equations in the system satisfy the adequacy conditions of over-identification.

Another characteristic of the model is that GDP growth is not considered endogenous because the share of regional banks in China to the total loan is low compared with national and joint-equity banks. Therefore, it is reasonable to assume that only the economic environment will influence the loan growth, CAR, and NPLs of the regional banks, whereas loan growth of the regional banks will not affect the regional economic environment, thereby simplifying the model.

7.2.2 Data used in the analysis

One of the problems in empirical research of banking system in China is the sample size, particularly before 2008. For a long time, only about 20 large banks have available detailed data (5 large state-owned banks, 12 joint stock banks, and 3 city commercial banks). The quality of the data is also notoriously inaccurate and inconsistent with Western standards. Since 2006, however, things have significantly improved. As a key component of the financial reform policy, CBRC required all banks in China to publish their annual financial reports. Most choose to publish their reports in the newspaper "king rong shi bao" (Financial News). Others choose to publish them in regional newspapers or on their website homepages. The quality of data also notably improved. These improvements greatly facilitate the researches in the field.

Data used in this chapter are obtained partly from the annual reports of Chinese regional banks published in the newspaper "king rong shi bao," and partly from the websites of the banks. Data on regional economies come from the yearbooks of statistics published in the website of local governments. The time spread is from 2005 to 2008.

The number of banks with data available varies over the years. The earliest and the latest annual data available are not the same for each bank because of the time of establishment or the preference of banks. Thus, our database is an unbalanced panel database. Due to the design of the model, a bank is included in the sample of that year only when the data in a year and the year before that year (its AR(1) term) are both available I also delete some outliers. After these considerations, the final sample includes 114 banks in 2008, 101 in 2007, 80 in 2006. Among these banks, 12 are rural commercial banks. However, these banks also operate in the middle and large cities. Table 7.1 shows the descriptive statistics of variables used in the model.

Table 7.1. Descriptive statistics of the variables

Unit: 100 million RMB Yuan, %

	Asset	Capital	Car	Deposit	Foreign	GDP	Gov	Loan	ni	npl
2008										
Mean	404.72	28.99	12.46	325.17	2.69	2226.85	21.06	207.99	1.03	2.23
Median	173.65	10.66	11.34	148.62	0.00	1400.91	17.30	101.95	1.02	1.97
Maximum	4170.21	480.94	38.09	3158.40	70.00	13698.15	91.33	1930.74	2.88	9.80
Minimum	26.03	1.91	8.11	23.26	0.00	233.90	0.00	7.25	0.02	0.18
Std. Dev.	638.49	62.03	3.92	495.89	8.70	2557.51	18.14	316.02	0.51	1.44
2007										
Mean	331.47	18.28	10.71	262.32	2.30	1880.99	21.25	169.76	0.90	2.50
Median	142.31	8.37	10.13	112.30	0.00	1150.90	18.69	78.12	0.88	2.13
Maximum	3542.23	266.68	30.67	2596.87	70.00	12188.85	91.33	1572.08	6.46	16.56
Minimum	21.76	0.73	2.76	19.74	0.00	193.50	0.00	6.27	-0.51	0.00
Std. Dev.	534.78	33.61	3.72	410.97	8.31	2245.84	18.06	257.60	0.72	2.04
2006										
Mean	280.66	13.69	9.39	232.64	2.54	1662.43	22.14	149.57	0.62	4.08
Median	136.11	7.45	8.96	96.02	0.00	956.09	18.58	83.13	0.56	3.32
Maximum	2729.69	123.41	62.04	2330.90	90.00	10366.37	91.33	1295.77	2.89	21.59
Minimum	17.54	0.77	-0.09	15.27	0.00	158.60	0.00	8.89	-1.39	0.00
Std. Dev.	447.40	20.69	6.15	382.13	10.33	1989.66	19.02	222.30	0.49	3.48
2005										
Mean	250.20	11.03	7.64	209.37	2.63	1565.73	20.09	131.75	0.76	5.81
Median	136.43	5.39	8.11	115.27	0.00	1112.40	16.35	80.03	0.46	4.86
Maximum	2401.36	104.29	23.37	1985.47	90.00	9164.10	72.73	1221.41	5.96	22.79
Minimum	13.63	0.70	-3.90	11.86	0.00	142.60	0.00	6.56	0.01	0.00
Std. Dev.	418.08	17.61	4.09	352.21	11.32	1856.84	16.82	205.72	0.89	4.52

Table 7.1 shows that the conditions of Chinese regional banks have significantly

improved after 2006. In 2005, average CAR of regional banks was 7.64%, slightly below the minimum requirement of the Basel Accord. Through the sample period, with stable speed (roughly two points each year) CAR gradually increased to 12.46%, which was significantly above the minimum level. In 2005, 37 banks (approximately 46% of the total sample banks in 2006) have CAR less than 8%. 64 (approximately 80% of the sample banks) banks' CAR was larger than 8%, but less than 10% (the level now considered adequate by Chinese supervisors for regional banks). In 2006, the number of banks with CAR was less than 8% was drastically reduced to 27 banks (approximately 26% of the total sample banks in 2007). The number of sample banks with CAR larger than 8%, but less than 10% declined to 73 (approximately 71% of the total sample banks). In 2007, the number of banks with CAR less than 8% was further reduced to 11 (approximately 10% of the total sample banks in 2008). The number of sample banks with CAR larger than 8%, but less than 10% was 53 (approximately 46% of the total sample banks). By 2008, all Chinese regional banks have met the minimum requirement (the minimum was 8.11%). Only 18 banks (approximately 16 % of the sample in 2008) did not meet the minimum 10% requirement for regional banks. Similarly, the net profit rate of the sample banks on average increased from 0.62% to 1% during the sample period. On average, NPL was almost reduced by half.

No large fluctuation in loan growth is found in the sample period. Nominal bank loan growth rate of sample banks on average declined slightly from approximately 24.1% in 2006 to 21.7% in 2007. Loan growth rate increased slightly to 22.8% in 2008. This was because in 2007 the government was worried about the overheating of the economy in 2007 and attempted to control loan growth. However, the subprime debt crisis broke out in 2008. The Chinese government planned a stimulus program worth RMB 4000 billion to counteract its negative effects. This stimulus program encouraged banks to offer more loans and increased the level of loan growth rate.

For GDP growth, on average the nominal GDP growth rate of operating areas for the sample regional banks has increased significantly from 16.24% in 2006 to 19.03% in 2007. The growth rate declined to 18% in 2008 because of the influence of the subprime crisis. This result is due to the large scale economic stimulus plan of the government.

7.3 Regression results

7.3.1 Results of OLS regression

For purposes of comparison, I first estimate the model using OLS approach equation by equation for each year. Table 7.2 presents the results of three equations using OLS.

Table 7. 2: Regression results of OLS

	dlnloan	*npl*	*car*
2008			
(Intercept)	-9.8524	2.4695***	5.4264***
	(-1.4426)	(3.9385)	(3.3996)
car08	1.3509***	-0.0346	
	(4.6756)	(-1.1906)	
car07			0.6759***
			(8.4913)
dlnasset			-0.1142***
			(-4.3176)
dlncapital			0.0564***
			(8.6458)
dlndeposit	0.6908***		
	(9.8431)		
dlngdp	-0.0450	0.0243	0.0245
	(-0.1762)	(1.1042)	(0.4275)
dlnloan		-0.0098	0.0556**
		(-1.4521)	(2.3719)
foreign	-0.1444	-0.0145	-0.0524*
	(-1.0948)	(-1.1504)	(-1.6603)
gov	0.0343	0.0027	-0.0245*
	0.5673	(0.4590)	(-1.7672)
ni		-0.0682	-0.5092
		(-0.3053)	(-0.9647)
npl08	-0.7510		-0.1976
	(- 0.9416)		(-1.0612)

npl07		0.4111***	
		(7.4288)	
Adjusted R-squared	0.5219	0.4196	0.5610
SD	17.2721	1.4551	3.9055
2007			
(Intercept)	-6.0289	2.0322***	5.2317***
	(-1.0285)	(2.6031)	(4.5296)
car07	0.7883**	-0.0705	
	(2.2160)	(-1.4322)	
car06			0.4598***
			(8.0304)
Dlnasset			-0.0338
			(-1.5292)
dlncapital			0.0483***
			(6.4291)
dlndeposit	0.8210***		
	(9.5183)		
dlngdp	-0.1778	0.0243	0.0045
	(-1.0976)	(1.1042)	（0.1343）
dlnloan		-0.0018	-0.0226
		(-0.1720)	（-1.2576）
foreign	-0.0041	-0.0043	0.0521
	(-0.0265)	(-0.2028)	（1.6187）
gov	0.0622	-0.0088	-0.0133
	(0.8669	(-0.8969)	（-0.8729）
ni		-0.2339	1.0336***
		(-0.9741)	（2.8895）
npl07	0.4853		0.0718
	(0.8005)		（0.5444）
npl06		0.2907***	
		(6.9841)	
Adjusted R-squared	0.4946	0.3862	0.5398
SD	17.2721	2.1451	3.7144
2006			
(Intercept)	1.0753	4.5630***	5.4926***
	(0.1002)	(2.5487)	（2.6864）
car06	1.8856***	-0.0500	
	(3.3780)	(-0.4763)	
car05			0.3338***
			（3.3814）
dlnasset			-0.0450**
			（-1.8842）
dlncapital			0.0252***

			(2.6038)
dlndeposit	0.3028***		
	(3.8639)		
dlngdp	0.1919	-0.0901	-0.0112
	(0.3878)	(-1.0487)	(-0.1228)
dlnloan		-0.0229	0.0680***
		(-1.2429)	(2.7798)
foreign	-0.3197**	0.0030	0.0711***
	(-2.4813)	(0.1302)	(2.8727)
gov	0.0626	-0.0185	-0.0034
	(0.8116)	(-1.3443)	(-0.2271)
ni		-0.7933	0.2062
		(-1.4246)	(0.3473)
npl06	-0.7184		-0.1002
	(-1.4914)		(-1.0261)
npl05		0.4364***	
		(7.5291)	
Adjusted R-squared	0.3302	0.5814	0.4864
SD	13.6934	3.0126	2.8344

***: significant at 1% level.

**: significant at 5% level.

*: significant at 10 % level

The first column shows the results for the first equation. The coefficient of variable *car* for three years is positive and significant. The result indicates that during the sample period, banks with high current CAR ratio tend to have high growth rate of loans. This result is consistent with the prediction of the "capital crunch" model. A comparison of the coefficient among the three years shows no clear breakdown in the structure during the sample period. The value of the coefficient is much lower in 2007 than the other two years. However, *t* value did not change over the period (all significant at 5% level).

For another endogenous variable *npl*, its sign is not consistent in the sample period. *npl* is negative for 2008 and 2006, but positive for 2007. However, *npl* is insignificant in all equations. This result is disappointing and may be due to the time lag of the effects of NPL as mentioned previously.

For the exogenous variables in column 1 of Table 7.2, we find that the coefficient of deposit growth rate is highly significant for every sample year. The coefficients are all positive, implying that banks with higher deposit growth rate will have higher loan growth rate. The results strongly imply that the deposit multiplier model may be a better explanation than the "capital crunch" model for loan growth in China during the sample period. The structure has a clear structural change. The value of the coefficient significantly increased in 2006, and its *t* value also gradually increased over the sample years.

The sign of the coefficient of *foreign* is negative in all three years, indicating that banks with a higher portion of foreign ownership tend to have lower loan growth rate. By contrast, the sign of the coefficient of *gov* is positive in the sample period. These are all what we expected. The sign of coefficient of *dlngdp* is negative in the years 2008 and 2007, which is not what we expected. Only in 2006 its sign is positive, suggesting that some counter-cyclical tendencies exist in loan growth. When the economy is in the period of overheating, bank will decelerate loan growth. However, all these exogenous variables are not significant, except for *dlngdp* in 2007 (weakly significant at 10% level) and *foreign* in 2006 (significant at 5% level).

7.3.2 Results of the model specification test

Several specification tests are conducted before the SEM is estimated using a systematic approach. The tests suggested by Hausman are used to test the specification.

7.3.2.1 Test for endogeneity

First, we need to ensure that the assumed endogenous variables in each equation are actually endogenous. Otherwise, using the SEM approach would be meaningless. Hausman (1978) suggests a simple test for endogeneity in a single equation. For multi-endogenous variables, we first run the reduced form regression by OLS for each endogenous variable and obtain their residuals. Then, we estimate a regression for each equation (extended by the residuals of the independent endogenous variables) in the system by OLS, assuming that all endogenous independent variables in each equation are exogenous. The first equation of the system is expressed as

$$d\ln loan_{it} = b_{01} + b_{11}\ car_{it} + b_{21} npl_{it} + b_{31} d\ln deposit_{it} + b_{41} d\ln gdp_{it} + b_{51} gov_{it} + b_{61} foreign_{it} + \rho_{11} v_{it2} + \rho_{21} v_{it3} + \varepsilon_{it1} \quad (7.2)$$

where v_{it2} is the residual of the reduced form regression for *npl* and v_{it3} is the residual of the reduced form regression for *car*. By construction, $\rho_{11} = 0$, if and only if *npl* is exogenous, and $\rho_{21} = 0$, if and only if *car* is exogenous. Thus, we can test their endogeneity using *t* test. We can also test the joint hypothesis H_0 : $\rho_{11} = 0$ and $\rho_{21} = 0$ by using the heteroskedasticity-robust Wald test. Similar procedures can be used on the two other equations in the system.

Table 7.3 shows the *t* value of the residuals of the reduced form equations. In the equation for loan growth, only *npl* is endogenous in the equation for 2007 (significant at 1% level). In the equation for *npl*, only *dlnloan* is endogenous in the equation for 2007 (significant at 5% level) and 2006 (weakly significant at 10% level). In the equation for *car*, *dlnloan* is endogenous in the equation for 2008 (significant at 5% level).

Table 7.3: Results of the endogeneity test, t statistics

	2008	2007	2006
dlnloan			
v_{it2}	0.2454	-1.7719***	-1.4020
	(0.1483)	(-2.4778)	(-0.8787)
v_{it3}	-0.4960	1.8368	1.0525
	(-0.8454)	(1.4093)	(0.9086)
npl			
v_{it1}	0.0030	0.0343**	0.0617*
	(0.2183)	(2.3724)	(1.6344)
v_{it3}	-0.0878	0.0276	-0.0397
	(-1.5081)	(0.3794)	(-0.1383)
car			
v_{it1}	0.3902**	0.0600	-0.0118
	(2.4422)	(1.5618)	(-0.0779)
v_{it2}	-0.5912	0.3301	0.0663
	(-1.5256)	(1.1741)	(0.2993)

Note: v_{it1} is the residual of the reduced regression for the first endogenous variable *dlnloan*;

v_{it2} is the residual of the reduced regression for the second endogenous variable *npl*; and

v_{it3} is the residual of the reduced regression for the third endogenous variable car.

Table 7.4 shows the results of the Wald test. In the system for 2008, only the equation for *car* passed the test, indicating that both endogenous variables are endogenous (significant at 5% level). The equation for *dlnloan* and *npl* equations for 2007 passed the test (all significant at 5% level). No equation passed the test for 2006.

Table 7. 4: Results of the endogeneity test-Wald test

	dlnloan		npl		car	
Value	Probability	Value	Probability	Value	Probability	Value
2008						
F-statistic	0.4014	0.6704	1.1562	0.3187	4.675484	0.0114***
2007						
F-statistic	4.3431**	0.0157	3.1005**	0.0498	1.928792	0.1513
2006						
F-statistic	1.9872	0.1446	1.5001	0.2302	0.068132	0.9342

In conclusion, I find that in each sample year at least one equation has an endogenous problem. But only in the equation for 2007，it is highly meaningful to use SEM approach to estimate the system, especially for the first equation.

7.3.2.2 Test for over-identification and application of 3SLS estimators

There are two major approaches to estimate a SEM: two-stage least square (2SLS) and three-stage least square (3SLS) approach.

2SLS approach used in SEM is the multi-equation version of 2SLS[10] approach used in a single endogenous equation in which each endogenous variable has

[10] Using 2SLS, we first regress each endogenous variable on all of instrumental variables for that endogenous variable and obtain its residuals, then again use these residuals as instruments to estimate the original model.

multi-instruments. We just apply 2SLS approach to estimate each equation in SEM.

2SLS approach does not consider possible correlations among the residuals of different equations. 3SLS approach overcomes this problem. In 3SLS, in first stage we estimate the model by 2SLS; in second stage we use its residuals to form a cross-equation covariance matrix ; in third stage, we use this covariance matrix to estimate the original model using FGLS method. This is the 2SLS version of the seemingly unrelated (SUR) model.

If all exogenous variables are uncorrelated with all residuals (which means the model is over-identified), both 2SLS and 3SLS estimators are consistent, but only 3SLS estimator is (asymptotically) efficient. Therefore, 3SLS is a better approach. On the other hand, if in each equation there are only enough exogenous variables in other equations which are uncorrelated with the error term in that equation (which means the model is just identified), 2SLS estimator is consistent, but 3SLS estimator is not, then we should choose 2SLS approach. Using this fact, we can use a Hausman test to check whether a model is over-identified and 3SLS approach is optimal. The null hypothesis is that all exogenous variables are uncorrelated with all disturbance terms. Let b_2 and b_3 be the vectors of coefficients estimated by 2SLS and 3SLS, respectively; under the null hypothesis, $b_2 - b_3 = 0$. Hausman test for over-identification is:

$$m = (b_2 - b_3)'(V_2 - V_3)^{-1}(b_2 - b_3)$$

where V_2 and V_3 are the estimated covariance matrices using 2SLS and 3SLS, respectively. Under the null hypothesis, the statistics follows F distribution. We should reject the null hypothesis if we obtain a high value of the statistics. Table 7.5 shows the results of the Hausman specification test for over-identification.

Table 7. 5: Test for over-identification

Year	statistic	Df	p-value
2008	17.3223	25	0.8698
2007	2.2524	25	1
2006	18.661	25	0.8131

P-values are all very high. Thus, we cannot reject the hypothesis that all exogenous variables are uncorrelated with all error terms, and that the three-stage approach is the most efficient estimator.

7.3.3 Results of SEM

Table 7.6 lists the results of SEM:

Table 7. 6: Regression results of SEM

	Dlnloan	npl	car
2008			
(Intercept)	-11.3093	2.2011 ***	3.3926
	(-1.3465)	(3.2693)	(0.8587)
car08	1.5631 ***	0.0044	
	(4.0626)	(0.1103)	
car07			0.8029 ***
			(4.3255)
dlnasset			-0.2551
			(-1.1801)
dlncapital			0.0385
			(1.4172)
dlndeposit	0.6888 ***		
	(9.3558)		
dlngdp	-0.0784	-0.0320	0.0423
	(-0.3003)	(-1.2867)	(0.3760)
dlnloan		-0.0136	0.1963
		(-1.5280)	(0.7089)
foreign08	-0.1696	-0.0174	-0.0685
	(-1.2378)	(-1.3458)	(-1.0957)
gov08	0.0371	0.0028	-0.0168
	(0.6109)	(0.4634)	(-0.5988)
ni		-0.1409	-0.2370

		(-0.6162)	(-0.2545)
npl08	-0.9774		0.2263
	(-0.7551)		(0.3622)
npl07		0.3993 ***	
		(7.0412)	
Adj R^2:	0.5186	0.4086	0.3524
SD:	11.5344	1.1233	3.1429
2007			
(Intercept)	-11.0828	0.1373	7.2720 ***
	(-1.0565)	(0.1323)	(3.8049)
car07	1.7358 ***	0.0342	
	(3.0579)	(0.5708)	
car06			0.4786 ***
			(6.7380)
dlnasset			-0.0245
			(-0.9319)
dlncapital			0.0477***
			(5.6445)
dlndeposit	0.8079 ***		
	(8.7555)		
dlngdp	-0.3519	0.0166	-0.0418
	(-0.9345)	(0.4509)	(-0.5138)
dlnloan		0.0151	-0.0611*
		(1.3651)	(-1.7728)
foreign07	-0.1463	0.0065	0.0472
	(-0.8445)	(0.3814)	(1.3192)
gov07	0.0140	0.0000	-0.0137
	(0.1858)	(0.0015)	(-0.8127)
ni		0.0083	0.8033 **
		(0.0469)	(1.9884)
npl07	0.2166		-0.0949
	(0.2717)		(-0.5189)
npl06		0.5040 ***	
		(12.0946)	
Adj R^2:	0.4503	0.6479	0.4581
Root MSE:	12.7727	1.2703	2.7742
2006			
(Intercept)	-2.5687	5.2820 *	3.5057
	(-0.1249)	(1.8824)	(1.4274)
car06	2.7549*	0.0646	
	(1.8031)	(0.2384)	
car05			0.4739 **
			(3.0029)
dlnasset			0.0811

			(0.9910)
dlncapital			0.0336*
			(2.7074)
dlndeposit	0.2560 ***		
	(3.2292)		
dlngdp	0.0535	-0.0852	0.0638
	(0.1028)	(-0.9705)	(0.6208)
dlnloan		-0.0925***	-0.0815
		(-3.2964)	(-0.5749)
foreign06	-0.4279 **	-0.0174	0.0442
	(-2.1501)	(-0.4799)	(1.2394)
gov06	0.0701	-0.0110	0.0137
	0.8555	(-0.7662)	(0.6553)
ni		-0.5488	0.1332
		(-0.9915)	(0.2474)
npl06	-0.9889		-0.0995
	-1.0201		(-0.5668)
npl05		0.3822***	
		(5.2941)	
Adj R^2:	0.2936	0.4979	0.1823
SD:	11.5090	2.1349	2.5631

***: significant at 1% level.

**: significant at 5% level.

*: significant at 10 % level

Column 1 of Table 7.6 shows the results of the first equation using the three-stage approach. Compared with OLS results, the results show that for the variable *car*, there is not much difference for the equation in 2008 and 2006. This is not surprising. For we find no serious endogenous problem in these two equations. By contrast, the coefficient of *car* more than doubled in the equation for 2007. Because of this substantial increase of coefficient in 2007, now we find a clear decline of the coefficient (which shows its economic significance) over the three years. The results indicate that the increase in loan growth caused by a unit increase of *car* gradually declines over the sample period. Correspondingly, we also find that *t* value of the coefficient (its statistical significance) improves over the period. In sum, we have a much clearer trend pattern on the relationship between loan growth and CAR than that in OLS estimates. This structural change is particularly obvious from 2006 to 2007. Therefore, we can conclude that the constraint of CAR on loan growth becomes increasingly tougher during the examined period.

The other endogenous variable *npl* remains insignificant for all sample years. However, its coefficient is reduced by more than half in the equation for 2007 compared with the results of OLS model.

The coefficients of exogenous variables in the equation for 2007 also have large changes compared with that of OLS estimator, but their significance did not change substantially.

Normally, using SEM will result in considerable loss of total fitness because of the partial correlation between instrumental and endogenous variables. However, in the SEM of this chapter we find no significant loss of fitness compared with the OLS model. Only a slight decline in the adjusted R squares in each year is observed,. This indicates that the instrumental variables chosen in the model have strong correlation with endogenous variables. Because of the insignificant changes in total fitness, the SEM also shows a trend of increasing total fitness over the sample years similar to OLS regressions.

Given that the two other equations are also identified, I also estimate them and list the results in Table 7.6. In the equation for *npl*, the results show that its AR(-1) term is positive and highly significant for each sample year. No other explanatory variable, except for *dlnloan* in the equation for 2006, is significant. Compared with the OLS results, the results are similar for 2008 and 2006, except that in OLS results. *dlnloan* in the equation for 2006 is insignificant. The coefficient of the AR term almost doubled in SEM results in the equation for 2007. The equation for *car* also shows a similar strong AR(-1) process, but the significance of the other explanatory variables is not as stable across the sample years as that in the other two equations. In the equation for 2008, the variables are all insignificant. In the equation for 2007, the coefficients of *dlnloan*, net income rate (*ni*), and capital growth rate (*dlncapital*) are significant, whereas only *dlncapital* is significant in the equation for 2006.

7.4 Conclusion

When CAR is implemented in a banking system that is generally under-capitalized, it is predicted that, if other factors remain the same, bank loan growth will be decreased. This section attempts to test this hypothesis by using a sample of regional banks in China from 2005 to 2008. I find that during the sample years, capital condition indeed became a strict constraint for loan growth of these banks. This constraint also became more stringent and significant across the sample years as the implementation of the policy continued.

The results of the endogenous tests show some endogenous relationship among loan growth, CAR ratio, and NPL ratio, particularly for 2007. As a result, the outcomes for 2007 using SEM estimator are not significantly different from that using OLS estimator. However, the coefficient of NPL is not significant either in the equation for loan growth or CAR for all three sample years. Further improvement for modeling of the behavior of NPL or data quality may produce better results.

Chapter 8 Conclusion

8.1 Summary of the dissertation

This dissertation focuses on the analysis of the effect of financial restructuring on regional banks. The dissertation is divided into two parts: the first part pertains to the effects of financial liberalization on the efficiency of regional banks whereas the second part deals with the influence of the reform of bank capital supervision on the loan supplies of regional banks. These two financial policies are normally the key items in restructuring programs of many countries. They reflect the two contradictory, but also complementary sides of any risk investment offered by a financial institution, namely, profit and risk.

For the first aim, the dissertation uses Japanese Shinkin banks as the object of analysis. The dissertation attempts to show how the productivity of Shinkin banks has changed during the period of FY 2001–FY 2008, and how this change is related to the bank restructuring during the same period. The dissertation focuses on the effects of M&A on efficiency and productivity changes. Encouraging M&A among financial institutions is an important policy in many financial liberalization programs. Governments hope that M&A will increase the efficiency and safety of these banks.

Chapter 4 estimates and examines the efficiency and productivity changes in Shinkin banks from FY 2001 to FY 2008. The dissertation attempts to obtain a more robust estimate by using a non-parametric bootstrapping technique. The dissertation finds that on average, the efficiency of Shinkin banks under the assumption of VRS declined significantly from FY 2005 to FY 2008, but the scale economy significantly increased in the same period. In both sample years, banks with medium scale assets are least efficient measured in the assumption of VRS, but they are most efficient in the scale economy. They are also most efficient when measured in the assumption of

CRS. This result is consistent with the theory of scale economy. From FY 2005 to FY 2008, the efficiencies of banks with small scale of assets have declined significantly declined, regardless of whether they are measured under the assumption of VRS or CRS. However, the scale economies of medium banks have increased significantly increased.

No significant changes in productivity are observed in FY 2001 to FY 2004. However, productivity significantly declined in FY 2005 to FY 2008. A major source of this trend is the deterioration in technical efficiency (the inward shift of the production frontier). This finding is consistent with the deterioration of the economic environment in the latter half of the 2000s. However, the scale economy of Shinkin banks improved in the second period, which partly offset the deterioration of the environments. This condition may have originated from the time lag in the effects of active M&A in the early 2000s.

Grouping banks according to the level of competition reveals the relationship between market power and productivity changes. From the original scores, the dissertation shows that banks located in the least competitive areas experience the least declines in productivity, but this result is not robust. Checking the components, we observe that banks located in the least competitive areas experience the slowest declines in pure efficiency. However, banks located in highly competitive areas are more successful in their efforts to decelerate the decline in technical efficiency. This deceleration compensates for their weakness in pure efficiency and the results are less clear.

Based on the estimates obtained from the analysis in Chapter 4, Chapter 5 investigates the effects of M&A activities on the efficiency and productivity changes in Shinkin banks. Merger incidents have no effects on the Malmquist index. Nevertheless it has significant effects on the efficiency scores and two components (technical and scale efficiency) of the Malmquist index.

The research also shows that some other factors closely related to M&A have significant effects on the efficiency and productivity changes. The variable of bank scale, indicators of organizational efficiency, and market share are significant at varying degrees in the model for efficiency scores. In the model for productivity changes, the indicator of change in bank scale is significant for all three equations, and the indicator for change of organizational efficiency is significant for the equation on Malmquist scores.

Therefore from the results we may say that on the whole the M&A activities that occurred during the early years of 2000s have significant positive effects on the efficiency and productivity changes of Shinkin banks. The results are consistent with general literature on the effects of M&A activities on the productivities of financial institutions.

The second part of the dissertation attempts to answer the question of whether the capital supervisory policy has significant effects on the risk taking behavior of regional banks. Chapter 6 first summarizes existing literature concerning the complex relationship among bank loan supply, capital position, and non-performing loans. Based on the literature, I develop a simple dynamic model to describe the interactions between these three variables. The model shows there is a maximum level of loan supply to retain NPL ratio at a minimum level. If the capital position of a bank cannot support this maximum level, its NPL level will be above the minimum level. This condition will reduce the profits of the bank and weaken its capital. Thus, a vicious cycle may be formed. The model shows that in this situation, strengthening the capital should be the first concern of banks.

Chapter 7 uses a sample of city commercial banks in China from 2005–2008 to analyze the endogenous relationship among loan supply, bank capital, and bank NPLs. The chapter attempts to answer the question of whether the implementation of CAR requirement since 2005 has had significant effects on the loan supplies of city commercial banks in China. To achieve this aim, I build a system consisting of three endogenous equations to describe the simultaneous determination of loan growth, NPL ratio, and CAR. Except for the use of simple OLS approach to estimate the system, I use a three-stage SEM approach to estimate the model. Both approaches find that CAR has very significant and positive effects on loan growth. Thus, the dissertation finds that during the sample years, capital condition indeed became a strict constraint for the loan growth of these banks, and the results are quite robust.

The dissertation also finds some endogenous relationship among loan growth, CAR ratio, and NPL ratio. As a result, the outcomes for 2007 using SEM estimator are significantly different from those using OLS approach. Now the value of the coefficient of CAR declines over the three sample years. The *t* value of the coefficient also improved over the period. Thus, we now have a much clearer trend pattern of the effects of CAR on loan growth. Both approaches also show that the total fitness of the model increases over time. Therefore, we may conclude that the constraint of capital becomes more stringent and significant across the sample years as the implementation of the policy continues.

8.2 Policy implications

The dissertation offers some empirical evaluation on the consequences of two policies, namely, encouraging M&A among financial institutions and tightening capital supervision. Both are key components of financial restructuring programs in many countries. Thus, the findings of this dissertation not only enrich available literature on financial restructuring, but also offer some guidance for financial reforms in the future.

The conclusions summarized above show that M&A activities among Shinkin banks that occurred during the early years of 2000s have significant positive effects on the efficiency and some components of productivity change in Shinkin banks. The evidence offers some support for policies that encourage M&A activities among banks. However, the evidence is not robust. In chapter 4, we find that medium sized banks are ranked first in the scale economy. Similarly, in chapter 5, we also find the variable *asset* (asset size) follows a quadratic form in the model to determine efficiency. We also do not observe significant and positive effects of market share on productivity changes. This is also a finding which is against the policy encouraging M&A between large Shinkin banks. Therefore, we should attempt to find the optimal scale for each kind of bank, and only encourage M&A among small banks; while discourage those among large banks.

The dissertation also supports the hypothesis that capital regulation has significant influence on loan supply. This also has important policy implications. First, the influence of capital regulation should be of great concern in the design of monetary policy. The effects of monetary policy will be influenced by bank capital conditions because of the existence of capital supervision, as we have explained in chapter 6. Banks are major providers of credits in many economies. If the tightening or relaxing of money supply by central banks cannot be transmitted smoothly through the banking system by the change in interest rate indirectly (interest rate channel) or by the change in loan supply directly (credit channel), the effect of monetary policy will be reduced or exaggerated. For example, when a government increases the monetary supply to stimulate an economy in recession, the increase in money supply may remain in the banks as involuntary deposit reserve because banks do not have adequate capitals. Then the effects of monetary policy may be reduced. The opposite may occur during periods of economic booms, when central banks wish to tighten the money supply. In these occasions, adjusting the capital regulation or injecting capital into the bank system using public money may make the monetary policy more effective.

The results also prove the existence of an endogenous relationship among loan growth, CAR, and NPL ratio. This finding also has significant policy implications. As analyzed in chapter 6, under-capitalized banks may be unable to maintain their loan supply at the optimal level, and thus they will also be inefficient and have NPL ratios above the optimal level. Under-capitalized banks are unable to replenish their capital because of inadequate profits and high level of NPL, which could result in the formation of a vicious cycle. In this case, government financial support to recapitalize these banks and delete their NPL may be beneficial to society. The lessons from countries that have experienced a “credit crunch” (such as the United States in the early 1990s and Japan in later 1990s) have repeatedly confirmed this result.

8.3 Suggestions for further studies

This dissertation leaves some questions unanswered. First, distribution efficiency of inputs and outputs following a price change is an important part of efficiency and productivity changes for a firm. In addition, price fixing ability is a key factor that will affect the efficiency measured in value terms of firms with market power. However, in the analysis of efficiency and productivity changes of Shinkin banks, the distribution effects are left untouched because of the difficulty in obtaining information of the prices of inputs and outputs of Shinkin banks. We do not know how the prices of inputs and outputs of Shinkin banks have changed during the sample periods, how they reacted to price changes by changing their distribution of inputs and outputs, and how these reactions have affected their efficiency and productivity changes because of this weakness. To find a method to measure the prices of inputs and outputs accurately for a financial institution and incorporate the assessed prices into the model of efficiency estimation in the future is a worthwhile exercise. In this manner, we can further decompose the productivity changes measured in value terms. Then, we will understand what proportion of this change is due to technical progress, what is due to pure efficiency improvement, what is due to the change in output and input prices, and what is due to the improvement of distribution of inputs and outputs in response to price changes. Aided by these findings, we can gain a better understanding on the sources of productivity changes in financial institutions.

Second, the dissertation does not consider the risk in the estimation of efficiency. As mentioned in the introduction, high returns in investment normally imply high risk. A bank may seem to be efficient because its assets have high returns, but these assets may have high risk. Thus, estimation without considering the risk is a biased measure of efficiency for financial institutions. Some studies have incorporated the risk into the estimation of efficiency, but in an indirect manner. For example, some researchers have added non-performing loans as an undesired (negative output) into the model in a DEA analysis. However, we need to determine a method that will consider directly the possibility of default in the model.

Finally, we use a non-parametric bootstrapping approach to obtain new samples. We use these new samples to calculate the expectation, standard error, bias, and significance level of the estimates. In Chapter 5, we actually also generate new samples through a semi-parametric bootstrapping approach. We use these samples to obtain similar statistics as that in the case of non-parametric bootstrapping. Comparing the results obtained from these two different approaches and checking the robustness of the conclusions we have drawn from the dissertation may be interesting.

There is also some room for improvement in the analysis of the relationship among risk, bank loan supply, and bank capital. In the analysis of Chapter 7, the results for variable *npl* (NPL ratio) are not satisfactory. The coefficient of *npl* is significant neither in the equation for loan growth nor for CAR for all three sample years, nor in any explanatory variable except its lagged term is significant in the equation for *npl*. Further improvement in the modeling of the behavior of NPL or data quality may produce better results.

Some analysts have suspected that the reaction of banks to their capital position may not be linear. Under-capitalized banks may be affected severely by capital regulation, whereas well-capitalized banks may not feel the influences of the regulation at all. This result implies the existence of one or even more regime switches in the model of loan supply. Similar phenomenon may exist in the relationship between *npl* and GDP growth rate. Only when GDP growth rate declines below a certain low threshold or exceeds a certain high threshold will *npl* be significantly influenced by GDP growth rate. Therefore, introducing the regime switch approach into the model of loan growth or *npl* and verifying if regime switches occur in these two models may be worthwhile.

In chapter 6, we see that the relationship between loan supply and GDP growth may also be endogenous. However, the effect of loan supply from these banks on economic growth is small because city commercial banks in China only offer a small portion of loan supplies. To simplify the system, I do not treat GDP growth rate as an endogenous variable, and set an equation for variable *gdp* (GDP growth rate) in the model. Hence, using a sample of large banks whose loan supplies have significant influence on GDP growth and examining the endogenous relationship between these two variables could also be a fruitful exercise.

References:

Agénor, P. R., J. Aizenman and A. W. Hoffmaister (2004). "The credit crunch in East Asia: what can bank excess liquid assets tell us?" Journal of International Money and Finance **23**(1): 27-49.

Agénor, P. R. and L. A. Pereira da Silva (2012). "Cyclical effects of bank capital requirements with imperfect credit markets." Journal of Financial Stability **8**(1): 43-56.

Aigner, D. J., C. A. K. Lovell and P. Schmidt (1977). "Formulation and estimation of stochastic frontier production function models." Journal of Econometrics **6**(1): 21-27.

Al-Sharkas, A. A., M. K. Hassan and S. Lawrence (2008). "The impact of mergers and acquisitions on the efficiency of the US banking industry: further evidence." Journal of Business Finance & Accounting **35**(1-2): 50-70.

Altunbas, Y., M.H. Liu, P. Molyneux and R. Seth (2000). "Efficiency and risk in Japanese banking." Journal of Banking & Finance **24**(10): 1605-1628.

Andersen, H. (2011). "Procyclical implications of Basel II: Can the cyclicality of capital requirements be contained?" Journal of Financial Stability **7**(3): 138-154.

Anderson, C. W. and I. Terry L. Campbell (2000). "Restructuring the Japanese banking system: Has Japan gone far enough?" International Review of Financial Analysis **9**(2): 197–218.

Antão, P. and A. Lacerda (2011). "Capital requirements under the credit risk-based framework." Journal of Banking & Finance **35**(6): 1380-1390.

Ataullah, A. and H. Le (2006). "Economic reforms and bank efficiency in developing countries: The case of the Indian banking industry." Applied Financial Economics **16**: 653-663.

Berger, A. N. and D. B. Humphrey (1997). "Efficiency of financial institutions: International survey and direction for future research." European Journal of Operational Research **98**(2): 175-212

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

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

Berger, A. N. and G. F. Udell (2002). "Small business credit availability and relationship lending- the importance of bank." The Economic Journal **112**(477): F32-F53.

Berglof, E. and G. Roland (1995). "Bank restructuring and soft budgets constraints in financial transition.pdf." Journal of Japanese and International Economies **9**: 354-375.

Bernanke, B. S. and C. S. Lown (1991). " The credit crunch." Brookings Papers on

Economic Activity **1991**(2).
Brana, S. and D. Lahet (2009). "Capital requirement and financial crisis: The case of Japan and the 1997 Asian crisis." Japan and the World Economy **21**(1): 97-104.
Calen, P. and R. Rob (1999). "The impact of capital based regulation on bank risk taking." Journal of Financial Intermediation **8**: 317-352.
Caprio, G. J. and L. H. Summers (1993). Finance and its reform : beyond laissez-faire, World Bank working paper.
Casu, B. and P. Molyneux (2003). "A comparative study of efficiency in European banking." Applied Economics **35**(17): 1865-1876.
Caves, D. W., L. R. Christensen and W. E. Diewert (1982). "The economic theory of index numbers and the measurement of input, output, and productivity." Econometrica **50**(6): 1393-1414.
Chiuri, M. C., G. Ferri and G. Majnoni (2002). "The macroeconomic impact of bank capital requirements in the emerging economies-past evidence to assess the future." Journal of Banking & Finance **26**(5): 881-904.
Cooper, W. W., L. M. Seiford and K. Tone (2006). Introduction to data envelopment analysis and its uses, Springer.
Dai, j., y. Jin and B. Liu (2008). "Capital regulation, bank lending and asymmetric effects of monetary policy." China Economic Quaterly **7**(2): 481-509. In Chinese.
Debreu, G. (1951). "The coefficientof resource utilization." Econometrica **3**(19): 273-292.
Devaney, M. and W. L. Weber (2000). "Productivity growth, market structure, and technological change: evidence from the rural banking sector." Applied Financial Economics **10**: 587-595.
Diamond, D. W. (1984). "Financial intermediation and delegated monitoring." Review of Economic Studies **51**(3): 393-414.
Diamond, D. W. and P. H. Dybvig (1983). "Bank runs, deposit insurance,and liquidity." Journal of Political Economy **91**(3): 401-419
Dietsch, M. and A. Lozano-Vivas (2000). "How the environment determines banking effciency: A comparison between French and Spanish industries." Journal of Banking & Finance **24**: 985-1004.
Drake, L. and M. Hall (2003). "Efficiency in Japanese banking: An empirical analysis." Journal of Banking & Finance **27**(5): 891-917.
Drake, L., M. J. B. Hall and R. Simper (2009). "Bank modelling methodologies: A comparative non-parametric analysis of efficiency in the Japanese banking sector." Journal of International Financial Markets, Institutions and Money **19**(1): 1-15.
Farrell, M. J. (1957). "The measurement of productive efficiency." Journal of the Royal Statistical Society **serial A**: 253-281.
Fethi, M. D. and F. Pasiouras (2010). "Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey." European Journal of Operational Research **204**(2): 189-198.
Fried, H. O., C. A. K. Lovell and S. S.Schmidt, Eds. (2008). The measurement of

productive efficiency and productivity growth, Oxford University Press.
Fried, H. O., C. A. K. Lovell and S. Yaisawarng (1999). "The impact of mergers on credit union service provision." Journal of Banking & Finance **23**: 367-386.
Fukuyama, H. (1993). "Technical and scale efficiency of Japanese commerical banks: a non-parametric approach." Applied Economics **25**(8): 1101-1112.
Fukuyama, H. (1996). "Return to scale and efficiency of credit associations in Japan： A nonparametric frontier approach." Japan and World Economy **8**: 259-277.
Fukuyama, H. and W. Weber (2008). "Japanese banking inefficiency and shadow pricing." Mathematical and Computer Modelling **48**(11-12): 1854-1867.
Garcia-cestona, M. and J. Surroca (2008). "Multiple goals and ownership structure: Effects on the performance of Spanish savings banks." European Journal of Operational Research **187**(2): 582-599.
Garden, K. A. and D. E. Ralston (1999). "The x-efficiency and allocative efficiency effects of credit union mergers." Journal of International Financial Markets, Institutions and Money **9**: 285-301.
Gilbert and P. W. Wilson (1998). "Effects of Deregulation on the Productivity of Korean Banks." Journal of Economics and Business **50**: 133-155.
Grifell-Tatje, E. and C. A. K. Lovell (1996). "Deregulation and productivity decline:The case of Spanish savings banks." European Economic Review **40**: 1281-1301.
Hahn, F. R. (2007). "Domestic mergers in the Austrian banking sector: a performance analysis." Applied Financial Economics **17**(3): 185-196.
Hancock, D., J. A. Laing and A. J. Wilcox (1995). "bank capital shocks: dynamic effects on securities,loans and capital." Journal of Banking and Finance **19**: 661-677.
Harimaya, K. (2004). "Measuring the Efficiency in Japanese Credit Associations." Review of Monetary and Financial Studies **21**(92-112). In Japanese.
Hart, O. D. and D. M. Jaffee (1974). "On the application of portfolio theory to depository financial intermediaries." The Review of Economic Studies **41**(1): 129-147.
Hausman, J. A. (1978). "Specification test in econometrics." Econometrica **46**(6): 1251-1271.
Haynes, M. and S. Thompson (1999). "The productivity effects of bank mergers-evidence from the UK Building societies." Journal of Banking & Finance **23**: 825-846.
Heilpern, E., C. Haslam and T. Andersson (2009). "When it comes to the crunch: What are the drivers of the US banking crisis?" Accounting Forum **33**(2): 99-113.
Hellmann, T. F., K. C. Murdock and J. E. Stiglitz (2000). "Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?" The America Economic Review **90**: 147-165.
Horie, Y. (2010). "The operation area and the efficiency of the Shinkin banks." Economics Research **76**(7-40). In Japanese.
Horiuchi, A. and K. Shimizu (1998). "The deterioration of bank balance sheets in

Japan: Risk-taking and recapitalization." Pacific-Basin Finance Journal **6**(1-2): 1-26.
Hoshino, Y. (1992). Analysis on the merge among small and medium financial institutions, Taga Press. in Japanese.
Hyun, J.S. and B.K. Rhee (2011). "Bank capital regulation and credit supply." Journal of Banking & Finance **35**(2): 323-330. In Japanese.
Imai, M. and S. Takarabe (2011). "Transmission of liquidity shock to bank credit: Evidence from the deposit insurance reform in Japan." Journal of the Japanese and International Economies **25**(2): 143-156.
Isik, I. and M. K. Hassan (2003). "Financial deregulation and total factor productivity change: An empirical study of Turkish commercial banks." Journal of Banking & Finance **27**: 1455-1485.
Jacques, K. and P. Nigro (1997). "Risk-based capital, portfolio risk, and bank capital: A simultaneous equations approach." Journal of Economics and Business **49**(6): 533–547.
Jacques, K. T. (2008). "Capital shocks, bank asset allocation, and the revised Basel Accord." Review of Financial Economics **17**(2): 79-91.
Kahane, Y. (1977). "Capital adequacy and the regulation of the financial intermediaries." Journal of Banking and Finance **1**(2): 207-218.
Kao, C. and S. Liu (2009). "Stochastic data envelopment analysis in measuring the efficiency of Taiwan commercial banks." European Journal of Operational Research **196**(1): 312-322.
Kim, D. and A. M. Santomero (1988). "Risk in banking and capital regulation." The Journal of Finance **43**: 1219-1233.
Koehn, M. and A. M. Santomero (1980). "Regulation of bank capital and portfolio risk." The Journal of Finance **35**(5): 1235-1244.
Koopmans, T. C. (1951). Analysis of production as an efficient combination of activities. Activity analysis of production and allocation. T. C. Koopmans. New York, John Wiley & Sons. Monograph 13: 33-97.
Leland, H. E. and D. H. Pyle (1977). "Informational asymmetries, financial structure, and financial intermediation." The Journal of Finance **32**(2): 371-387.
Lin, Y. (2007). "Developing small and medium banks to impropve financial structure." China Center for Econmic Research, Working paper, in Chinese,.
Lown, C. S. and J. Wenninger (1992). "The role of banking system in the credit crunch." The Federal Reserve Banks of New York, Working paper
Makinnon, R. (1973). Money and capital in economic development. Washington D. C. U.S.A, Brookings Institution Press.
McKillop, D. G., j. C. Glass and Y. Morikawa (1996). "The composite cost function and efficiency in giant Japanese banks." Journal of Banking & Finance **20**(10): 1651-1671.
Meeusen, W. and J. v. D. Broeck (1977). "Efficiency estimation from Cobb-Douglas production functions with composed error " International Economic Review, **18**(2): 435-444
Meh, C. A. and K. Moran (2010). "The role of bank capital in the propagation of

shocks." Journal of Economic Dynamics and Control **34**(3): 555-576.
Montgomery, H. and S. Shimizutani (2009). "The effectiveness of bank recapitalization policies in Japan." Japan and the World Economy **21**(1): 1-25.
Mosser, P. C. and C. Steindel (1994). Economic activity and the recent slowdown in private sector borrowing. Studies on Causes and Consequences of the 1989-92 Credit Slowdown New York, Federal Reserve Bank of New York.
Otchere, I. and J. Chan (2003). "Intra-industry effects of bank privatization: A clinical analysis of the privatization of the Commonwealth Bank of Australia." Journal of Banking & Finance **27**(5): 949-975.
Paravisini, D. (2008). "Local bank financial constraints and firm access to external finance." The Journal of Finance **63**(5): 2161-2193.
Patrick, H. T. (1966). "Financial development and economic growth in underdeveloped countries." Economic Development and Cultural Change **14**(2): 174-189.
Peek, J. and E. S. Rosengren (2000). "Collateral damage: Effects of the Japanese bank crisis on real activities in the United States." American Economic Review **90**(1): 30-45.
Rebelo, J. and V. Mendes (2000). "Malmquist indices of productivity change in portuguese banking: The deregulation period." International Advances in Economic Research **6**(3): 531-543.
Pyle, D. H. (1971). "On the theory of financial intermediation." The Journal of Finance **26**: 741-747.
Rezitis, A. (2008). "Efficiency and productivity effects of bank mergers: Evidence from the Greek banking industry." Economic Modelling **25**(2): 236-254.
Rime, B. (2001). "Capital requirement and bank behavior: empirical evidence for Switzerland." Journal of Banking & Finance **25**: 789-805.
Rochet, J.C. (1992). "Capital requirements and the behaviour of commercial banks." European Economic Review **36**: 1137-1170.
Roy, P. V. (2004). The impact of the 1988 Basel accord on banks' capital ratios and credit risk-taking: an international study. EFMA 2004 Basel Meetings.
Saunders, A. (2002). "Comments on the macroeconomic impact of bank capital requirements in emerging economies." Journal of Banking & Finance **26**: 905–907.
Schumpeter, J. A. (1939). Business cycle, Cambridge Univ Press.
Shaw, E. S. (1973). Financial deepening in economic development. New York, Oxford University Press.
Shen, Y., M. Shen, Z. Xu and Y. Bai (2009). "Bank size and small- and medium-sized enterprise (SME) lending: evidence from China." World Development **37**(4): 800-811.
Shephard, R. W. (1953). Cost and production functions. Princeton, Princeton University Press.
Shrieves, R. E. and D. Dahl (1992). "The relationship between risk and capital in commercial banks " Journal of Banking & Finance **16**(2).
Simar, L. (1992). "Estimating efficiencies from frontier models with panel data: a

comparison of parametric, non-parametric and semi-parametric methods with bootstrapping." Journal of Productivity Analysis **3**: 171-203.

Simar, L. and P. Wilson (2007). "Estimation and inference in two-stage, semi-parametric models of production processes." Journal of Econometrics **136**(1): 31-64.

Simar, L. and P. l. W. Wilson (1998). "Sensitive analysis of efficiency scores: How to bootstrapping in Non-parametric frontier models." Management Science **44**(1): 49-61.

Simar, L. and P. W. Wilson (1999). "Bootstrapping the malmquist index." European Journal of Operational Research **115**: 459-471.

Stiglitz, J. E. and A. Weiss (1971). "Credit rationing in markets with imperfect information." The American Economic Review **71**(3): 393-410.

Syron, R. (1991). "Are we experience a credit crunch?" New England Economic Review: 2-10.

Tadesse, S. (2006). "Consolidation, scale economies and technological change in Japanese banking." Journal of International Financial Markets, Institutions and Money **16**(5): 425-445.

Vennet, R. V. (1996). "The effect of mergers and acquisitions on the efficiency and profitability of EC credit institutions." Journal of Banking & Finance **20**: 1531-1558.

Wang, O. and W. Wu (2012). "Capital supervision and bank credit growth." Economic Perspectives **3**: 63-66. In Chinese.

Watanabe, W. (2010). "Does a large loss of bank capital cause Evergreening? Evidence from Japan." Journal of the Japanese and International Economies **24**(1): 116-136.

Wheelock, D. C. and P. W. Wilson (2009). "Robust non-parametric quantile estimation of efficiency and productivity change in U.S. commercial banking." Journal of Business and Economic Statistics **27**(3): 354-368.

Wilson, P. (2008). "FEAR: A software package for frontier efficiency analysis with R." Socio-Economic Planning Sciences **42**(4): 247-254.

Wilson, P. W. (1993). "Detecing outliers in deterministic nonparametric frontier models." Journal of Business & Economic Statistics **11**(3): 319-323.

Woo, D. (2003). "In search of “Capital Crunch"- supply factors behind the credit slowdown in Japan." Journal of Money, Credit and Banking **35**(6): 1019-1038.

Wu, D. and J. Zhou (2006). "Capital requirement and commercial bank's operation: An empirical analysis of large and medium commercial banks in China." Journal of Financial Research **314**(8): 144-153. In Chinese.

Xu, W. and B. Chen (2009). "Bank lending and economic fluctuations in China : 1993-2005." China Economic Quarterly **8**(3): 969-994. In Chinese.

Xue, M. and P. T. Harker (1999). "Overcoming the inherent dependency of DEA efficiency scores: a bootstrap approach." Working paper, Wharton Financial Institutions Center, University of Pennsylvania.

Yang, Z. (2006). "A two-stage DEA model to evaluate the overall performance of Canadian life and health insurance companies." Mathematical and Computer

Modelling **43**(7-8): 910-919.

Zhao, X. and S. Wang (2007). "Impacts of capital regulation on credit growth of commercial banks:The Empirical study on Chinese banking sector between 1995 and 2003." Finance &Trade Economics **7**. In Chinese.