Stockholm School of Economics

Department of Finance

Tutor: Peter Englund

Cost Efficiency and Foreign Ownership in the U.S.: Stochastic Frontier Analysis

Edgaras Krusas*

December 2016

ABSTRACT

This study examines bank cost efficiency in the U.S. for the period 2008-2013. The main purpose is to investigate if foreign banks in the U.S. are less cost efficient than their domestic counterparts. I apply stochastic frontier approach which follows Battese and Coelli (1995) specification. Empirical results show significant associations of cost efficiency and bank-specific variables including leverage, size, output quality and performance indicators. The outcome of whether foreign banks are less cost efficient is inconclusive. I explain that heterogeneity in different bank groups can be a driving factor in inefficiency estimates thus more advanced methods are required in order to disentangle heterogeneity from inefficiency.

Keywords: banking, cost efficiency, foreign ownership, stochastic frontier analysis.

Acknowledgements: I would like to thank my supervisor Peter Englund for his support and feedback. I would also like to thank this course director Michael Halling for his comments regarding data collection and technical issues.

* 40936@student.hhs.se

Contents

1. Introduction

The banking industry in the U.S. undergone changing regulatory environment during 1980-2005 which had an impact on industry consolidation through mergers and acquisitions and an overall decrease in a number of commercial banks, see, for example, Montgomery (2003), Feng and Serletis (2009). The last two decades before the financial crisis in 2008-2009 also marked a period of economic integration which included increasing activities in international banking; international banks expanded funding, lending and capital markets operations in the U.S. via an increasing chain of branches and subsidiaries (Berlin, 2015). The decision between these two organizational forms lies in a combination of economic, political, regulatory and strategic factors although regulation and taxes appear to be the most significant ones. (Cerutti et al, 2007).

After the end of financial crisis, regulators concluded that international banking has grown outside the scope of national supervision; the supervisory focus shifted from efficiency benefits towards potential crisis costs, risk management and stress-testing. However, additional regulations towards foreign banks in the U.S. might not only create more financial stability but also impact operational efficiency. Thus, regulators, economists and policy-makers should be concerned about a potential what foreign banks will be put at disadvantage in comparison to their domestic counterparts and begin to shift their activities to other countries with more satisfactory regulatory environment, the process which is also known as capital flight (Berlin, 2015).

The thesis aims to contribute to the existing literature about foreign ownership and banking efficiency in the U.S. It is important to study whether foreign banks were operating below, above, or at the same cost efficiency as domestic banks during 2008-2013, just before the implementation phase and actual compliance dates of new regulations during 2014-2018 (Federal Reserve Press Release, 2014). If foreign banks were already less cost efficient before 2014-2018, such regulatory changes might further impact ability to compete in the U.S. banking sector resulting in an overall decrease of global banking operations. Thus, the research question proposed in this quantitative study is as follows: **Are foreign banks less cost efficient than domestic banks in the U.S.?**

In order to answer the research question, I will use stochastic frontier approach. This approach helps to tackle two major areas of interest in banking efficiency studies. First, it allows to estimate cost (or profit) efficiency of banks. Second, it might help to find exogenous variables describing the environment of banking activities which are significantly related to inefficiency and perhaps can explain at least some of the differences across banks in a given sample (Kumbhakar and Lovell, 2000).

The focus of this study is entirely on cost efficiency. According to Feng and Serletis (2009), majority of studies which used stochastic frontier approach in estimating banking efficiency in the U.S. chose cost efficiency approach. Also, cost efficiency is associated with greater financial intermediation and allows to benchmark banks in terms how well they are performing, managing their costs and achieving similar outputs for a given set of input prices in a current regulatory environment (Fries and Taci, 2005).

The thesis is structured in a following way. Literature review is provided in section 2. An overview of studies focusing on bank efficiency in the U.S. and elsewhere is provided in this section, with the main focus on research related to foreign ownership analysis. Methodology is presented in section 3. Stochastic frontier analysis is explained in detail with a particular attention to Battese and Coelli (1995) specification, its benefits and drawbacks. Data part is provided in section 4. Among other things, the choice of value-added approach related to taking a particular set of outputs and input prices for the model is explained in this section. Empirical results are provided in section 5. Finally, section 6 concludes.

2. Literature review

2.1. Bank efficiency studies

Bank efficiency studies in general apply different techniques in order to benchmark the best performing banks in terms of defined measure such as production, cost, revenue or profit. Different techniques yield quite disperse results thus it is important to understand all underlying factors in methodology specification and data quality. There is no best practice but some models might be preferred given certain circumstances which will be described in the following sections. In this section I will provide some details about the results on past research on bank efficiency focusing on the U.S. and other countries.

Berger and Humphrey (1997) provide a survey of over 120 studies which attempt to apply various frontier efficiency analysis techniques for financial institutions in 21 countries. Studies using parametric approaches tend to lower dispersion in mean efficiency estimates. Although industry efficiency estimates were found in line among all used approaches (for example, SFA efficiency estimates often vary between 0.7 and 0.9 based on the list of investigated studies), individual firm efficiency rankings were inconsistent thus it was concluded that analysis of firm specific efficiency correlates should be viewed with caution (Berger and Humphrey, 1997). Also, the authors suggested to use comparison of efficiency estimates in groups of observations instead of individual observations in future research and mentioned that there is a considerable lack of information regarding the main determinants of efficiency differences among firms (Berger and Humphrey, 1997).

Mester (1997) finds that differences in market environment where banks are operating might distort inefficiency scores if not controlled in a cost function. For example, when accounting for output quality and riskiness, the author managed to tackle an upward bias of inefficiency estimates for U.S. banks observed in previous studies. Also, Mester suggested that a single cost model for all groups is not sufficient to account for heterogeneity in the U.S. banking market as it results in an upward bias in inefficiency estimates (1997).

Berger and Mester (1997, 2003) attempt to compare cost, standard profit and alternative profit efficiency concepts among U.S. banks and find out that each concept adds an additional informational value. For example, the authors find that large banks (with assets over USD 10 billion) are 2.5% more cost efficient than small banks (with assets under USD 0.1 billion) but in terms of profit efficiency small banks show the best results (Berger and Mester, 1997). Among other variables, the authors account for equity capital and non-performing loans (1997, 2003). Among other findings, the authors mention that the cost productivity worsened while profit productivity improved under period of investigation (Berger and Mester, 2003).

Feng and Serletis (2009) attempt to use a globally flexible Fourier cost functional form in U.S. banking industry – the findings conclude that their method is able to tackle at least some of functional form misspecification prominent in less advanced techniques. The authors show that the largest subgroup of banks with assets exceeding 1 billion dollars (in 1998 terms) are the least cost efficient during 1998-2005. However, their findings also suggest than when excluding this largest subgroup of banks, larger banks exhibit higher cost efficiency level comparing to medium banks and medium banks are more cost efficient than the smallest banks in the sample (Feng and Serletis, 2009).

2.2. Bank efficiency and foreign ownership

This study attempts to uncover the role of foreign ownership on bank performance. The literature of similar studies between domestic and foreign banks is extensive and covers many different countries. Banks in both developed nations and transition economies are analyzed. The main focus of this literature review is on the U.S. market, however, knowledge of bank efficiency trends in other countries is beneficial to gain a broader understanding what determines better performance of foreign banks in different settings.

To begin with, Chang et al (1998) makes a cost efficiency comparison of foreign banks in the U.S. and their domestic counterparts by using a multiproduct translog stochastic-cost frontier model. Their results show that domestic banks are more cost efficient. An interesting addition to their findings is that large banks holding a higher share of foreign assets were less efficient during the 1990s (Chang et al, 1998).

Similarly, DeYoung and Nolle (1996) makes profit efficiency analysis of foreign versus domestic banks in the U.S. and finds identical outcome that foreign banks are less profit efficient. The reason of relatively higher inefficiency of foreign banks was reliance on expensive purchased funds, according to the authors.

Berger et al (2000) studies cross-border banking efficiency in France, Germany, Spain, the U.K., and the U.S. and finds that foreign banks are less cost efficient than domestic banks in both cost and profit efficiency. However, banks from the U.S. are found to be consistently outperforming domestic banks in cost efficiency terms. This is an important finding suggesting that some banks from developed countries can be more efficient when operating as foreign banks comparing to domestic banks in other developed countries (Berger et al, 2000).

Bonin et al (2005), and Fries and Taci (2005) studies cost efficiency of foreign versus local-based banks in 6 and 15 countries in eastern Europe, respectively. Their findings are the opposite to studies concentrated on developed markets – foreign-owned banks are significantly more cost efficient. State-owned banks are found to be the least cost efficient among all banks operating in the developing economies (Fries and Taci, 2015). Similarly, banks which are privatized in an earlier stage are more cost efficient than later-privatized financial institutions (Bonin et al, 2005).

In addition, Poghosyan and Borovicka (2007) studies 19 European emerging markets and instrument for the decision of foreign owners to acquire domestic banks. Their findings do not follow the general consensus as they identify that foreign owners participated in a “cream-skimming” effect and acquired most efficient banks in the first place. After accounting for this their results no longer show significance of foreign banks versus domestic counterparts and even change sign of the relationship of foreign ownership on cost efficiency.

Except of few exceptions like the one mentioned above, there is a general consensus view that foreign banks are in fact less cost (and profit) efficient in developed nations while the opposite is correct in countries under transition which lack well-developed institutions (see Berger (2007) who reviews over 100 different studies and Lensink et al (2008), Lensink and Meesters (2014) who investigate several thousand banks in over 100 different nations).

3. Methodology

3.1. Parametric versus non-parametric approaches

Both parametric and non-parametric techniques are employed in order to estimate inefficiency scores and to separate the best performing banks from others operating at a somehow sub-optimal performance level. It is important to note that neither parametric nor non-parametric technique is considered superior in general – there are certain conditions when one might be preferred to the other.

The main advantage of non-parametric methods such as data envelopment analysis (DEA) and free disposal hull (FDH) approach is that they do not impose a strict functional form for a frontier shape. However, the key weakness is a general assumption of no random error allowing for luck, data quality issues or measurement errors to influence efficiency estimates (Kumbhakar and Lovell, 2000).

Parametric techniques such as stochastic frontier analysis (SFA), distribution-free approach (DFA) and thick frontier approach (TFA) do not suffer from the abovementioned weakness relevant for non-parametric models thus potential errors or inaccurate estimates in accounting data do not deteriorate banking efficiency analysis using parametric methods. However, the key disadvantage of parametric techniques is imposition of a specific functional form, for example, Cobb-Douglas and translog, which in one case might be not flexible enough to approximate actual data, and in other case it might result in over-specification and multicollinearity issues.

Comparing three parametric models, SFA is preferred to the other two approaches because TFA does not provide cost efficiency estimates for each individual firm and discards at least half of available observations (the latter is often the most important drawback for researchers); DFA is based on assumption that cost efficiency is time-invariant – that can be considered as a very strong proposition specially in longer panels (Kumbhakar and Lovell, 2000).

3.2. Technical background of SFA

Stochastic frontier analysis (SFA) originated from two papers issued in 1977 (Aigner et al; Meeusen and van den Broeck). Both papers included a unique composed error structure of this parametric approach consisting of a traditional symmetric random-noise component and a one-sided inefficiency term. Since 1977 there were many attempts to benchmark different banks in terms of their production, cost and profit frontiers.

Early SFA models allowed to derive mean efficiency estimated over the sample but not by each individual observation. Jondrow et al (1982) were the first to propose a model extension which allowed to estimate individual technical inefficiencies (referred as JLMS estimates) within a given sample.

SFA panel data models were proved as a better alternative comparing to cross-sectional setting. Schmidt and Sickles (1984) noted three limitations which could be tackled by using panel data:

- Strong distributional assumptions towards estimation of stochastic frontier model and decomposition of inefficiency error term from statistical noise;
- Independence assumption that inefficiency error term is independent of the regressors (outputs such as loans and inputs such as labor or physical capital);
- Efficiency estimates are not consistent since conditional mean (or mode) of each individual observation does not achieve convergence to zero as the size of the cross section increases.

Availability of panel data can overcome each of these problems. Moreover, time-invariant efficiency was relaxed in panel data models in the beginning of 1990s by the papers of Cornwell et al (1990), Kumbhakar (1990), and Battese and Coelli (1992). At a similar timing, two-step approach of estimating determinants of efficiency variation was complemented by a more advanced one-step approach (see, for example, Kumbhakar et al, 1991; Huang and Liu, 1994; Battese and Coelli, 1995). Two-step approach (during which inefficiency is estimated in the first step, and then used as a dependent variable in a regression in the second step) was later criticized to incorporate substantial omitted variable bias by Wang and Schmidt (2002).

They provide evidence of the bias on the second stage parameters (based on Monte Carlo simulations) and propose that this bias could be avoided by a one-step approach which is estimated in a maximum likelihood setting (Wang and Schmidt, 2002).

Greene (2005a and 2005b) noted that previous approaches suffer from heterogeneity in inefficiency term and proposed “true” fixed and random-effects approaches which account for this heterogeneity and firm efficiency. Greene also noted that his models suffer from phenomena called incidental parameters problem (Neyman and Scott, 1948; Lancaster, 2000). In panel data setting, as the number of observations grow, unknown parameters increase at the same rate resulting in inconsistent efficiency estimates (Kumbhakar and Tsionas, 2011). In cross-sectional setting, incidental parameters problem is reflected by the third limitation noted by Schmidt and Sickles (1984) which is shown in the third bullet-point above.

Recently, some more advanced approaches appeared in SFA which seem to develop techniques which tackle long-existed deficiencies in previous methods. Wang and Ho (2010) developed a transformed fixed-effects approach which solves incidental parameters problem. Chen et al (2014) proposed a within maximum likelihood approach, which, based on their findings, also overcomes this same problem. Finally, Belotti et al (2013), Belotti and Ilardi (2015) proposed new models based on Chen et al (2014) and their own research.

Kumbhakar and Tsionas (2011) made an overview of some recent developments in the field of SFA including estimation of latent class models which address technological and behavioral heterogeneity and models using local maximum likelihood method (LML). The latter models tackle the limitations of SFA models comparing to non-parametric methods (for example, DEA). In particular, the most important limitation – functional form misspecification – is discarded by using LML. (Kumbhakar and Tsionas, 2011).

3.3. General form and error term specification

Stochastic frontier analysis can accommodate both cross-sectional and panel data. In this study I will focus entirely on panel data analysis. The general form of cost function in time-varying stochastic frontier model for panel data is as follows:

$$TC_{it} = c\,(Y_{it}, X_{it}) + \varepsilon_{it}$$

TC_{it} is a vector of total costs, Y_{it} is a vector of outputs, X_{it} is a vector of input prices, ε_{it} is an error term and $c\,(\cdot)$ is a functional form. Subscripts i and t account for an individual bank and time, respectively.

The error term ε_{it} in SFA can be separated into two different error components – random noise (v_{it}) and inefficiency term (u_{it}):

$$\varepsilon_{it} = v_{it} + u_{it}$$

In this regard v_{it} corresponds to random error which is identically and independently distributed (i.i.d.) and follows standard normal distribution by assumption:

$$v_{it} \sim N\,(0, \sigma_v^2)$$

u_{it} corresponds to one-sided inefficiency term which signals how far above a bank is from cost efficiency frontier and is also i.i.d. In addition, inefficiency component is non-negative as it measures deviations from efficiency which is assumed to be zero at this point (later in this section, I will include an additional specification):

$$u_{it} \sim N^{+}(0, \sigma_u^2)$$

Finally, the final distributional assumption states that v_{it} and u_{it} are independent of each other and of the regressors (Kumbhakar and Lovell, 2000).

The estimation procedure I will estimate and declare in the following sections will produce two parameters for error terms – sigma-squared (σ^2) and gamma (γ). The parameter σ^2 shows total sum of variance of both error components (random noise and inefficiency term):

$$\sigma^2 = \sigma_v^2 + \sigma_u^2$$

The parameter γ lies between zero and one and indicates importance of inefficiency component. In case this parameter equals zero, inefficiency term is irrelevant and all error refers to random noise. In case it equals one, all deviations from a frontier can be explained by inefficiency term:

$$\gamma = \frac{\sigma_u^2}{\sigma^2}$$

3.4. Model specification

There are several model specifications available for production and cost functions that differ in flexibility – Cobb-Douglas (where both dependent and independent variables are linear in logarithms), translog (where both dependent and independent variables are quadratic in logarithms), quadratic, Fourier (translog function augmented with trigonometric Fourier terms), etc. The choice of a right model depends on several aspects including sample size, choice of outputs and input prices, prior knowledge or expectation of a true functional form.

Flexible forms such as translog or Fourier have an advantage over strict imposition of a particular functional form on given economic variables what is considered as the most important problem in parametric models. This advantage comes at a cost of potential multicollinearity issue among regressors which is particularly severe in small samples and functional forms containing many parameters (including both outputs and input prices) (Kumbhakar and Lovell, 2000).

Cobb-Douglas specification can be more relevant (and sometimes the only choice) for smaller samples due to multicolinearity. In addition, the analysis of inefficiency correlates of a one-sided inefficiency term is more direct and straightforward in Cobb-Douglas case. On the other hand, Cobb-Douglas cannot accommodate multiple outputs without violation of obligatory curvature properties in output space; this drawback was noted long time ago by Hasenkamp (1976). Also, an assumption of over-simplistic functional form might be inaccurate both for a single output and multiple-outputs leading to biased estimates of both random error term and inefficiency term (Kumbhakar and Lovell, 2000).

I will compare several combinations of translog specification and I also use Cobb-Douglas specification in order to check robustness of the obtained results. Translog specification for K outputs and L input prices can be summarized in a following way:

$$\ln(\mathrm{TC}/\mathrm{X}_{2\mathrm{it}}) = \beta_0 + \sum_{k=1}^{K} \alpha_k \ln \mathrm{Y}_{\mathrm{kit}} + \frac{1}{2}\sum_{k=1}^{K} \sum_{m=1}^{K} \alpha_{km} \ln \mathrm{Y}_{\mathrm{kit}} \ln \mathrm{Y}_{\mathrm{mit}} +$$

$$\sum_{l=1}^{L} \beta_l \ln (\mathrm{X}_{\mathrm{lit}}/\mathrm{X}_{2\mathrm{it}}) + \frac{1}{2}\sum_{l=1}^{L} \sum_{n=1}^{L} \beta_{ln} \ln (\mathrm{X}_{\mathrm{lit}}/\mathrm{X}_{2\mathrm{it}}) \ln (\mathrm{X}_{\mathrm{nit}}/\mathrm{X}_{2\mathrm{it}}) +$$

$$\sum_{k=1}^{K} \sum_{l=1}^{L} \varphi_{kl} \ln \mathrm{Y}_{\mathrm{kit}} \ln (\mathrm{X}_{\mathrm{lit}}/\mathrm{X}_{2\mathrm{it}}) + \mathrm{v}_{\mathrm{it}} + \mathrm{u}_{\mathrm{it}}$$

TC are total costs, $\ln \mathrm{Y}_{\mathrm{kit}}$ are output quantities in natural logs, $\ln (\mathrm{X}_{\mathrm{lit}}/\mathrm{X}_{2\mathrm{it}})$ are input prices in natural logs. Division by one of input prices $\mathrm{X}_{2\mathrm{it}}$ imposes homogeneity of the cost function in prices (both dependent variable TC and all input prices are deflated by a chosen price, in this example $\mathrm{X}_{2\mathrm{it}}$). There is an additional requirement that symmetry restrictions on second partial derivatives $\alpha_{km} = \alpha_{mk}$ and $\beta_{ln} = \beta_{nl}$ would be imposed (Kumbhakar and Lovell, 2000). If $\mathrm{K} = 1$, equation collapses to a single-output translog cost frontier. If also $\beta_{ln} = \varphi_l = 0$ then this equation becomes linear-in-logarithms (nested) Cobb-Douglas cost frontier:

$$\ln(\mathrm{TC}/\mathrm{X}_{2\mathrm{it}}) = \beta_0 + \alpha_1 \ln \mathrm{Y}_{\mathrm{it}} + \sum_{l=1}^{L} \beta_l \ln (\mathrm{X}_{\mathrm{lit}}/\mathrm{X}_{2\mathrm{it}}) + \mathrm{v}_{\mathrm{it}} + \mathrm{u}_{\mathrm{it}}$$

3.5. Cost efficiency scores and inefficiency correlates

Cost efficiency scores can be measured over time after cost efficiency frontier is estimated given all available outputs and input prices. A stochastic cost frontier can be described in a bit different fashion comparing to (1) equation in section 3.3 as follows:

$$\mathrm{TC_{it}} \geq \mathrm{c}\ (\mathrm{Y_{it}}, \mathrm{X_{it}}) \cdot \exp\{\mathrm{v_{it}}\}$$

Again, $\mathrm{TC_{it}}$ reflects total costs / expenditure of each individual bank. The first part of cost frontier [$\mathrm{c}\ (\mathrm{Y_{it}}, \mathrm{X_{it}})$] is called a deterministic part which is common to all banks and the second part [$\exp\{\mathrm{v_{it}}\}$] is a random part which captures bank-specific random shocks.

Then, cost efficiency ($\mathrm{CE_{it}}$) can be estimated in the range [0; 1] in a following way:

$$\mathrm{CE_{it}} = \frac{\mathrm{c}\ (\mathrm{Y_{it}}, \mathrm{X_{it}}) \ \cdot \ \exp\{\mathrm{v_{it}}\}}{\mathrm{TC_{it}}} = \exp\{-\mathrm{u_{it}}\}$$

An individual bank is considered as cost efficient if its total costs given available outputs and input prices lay on an estimated cost frontier reflecting a benchmark of the best banks in terms of cost efficiency in a given sample. Alternatively, an individual bank has a certain degree of cost inefficiency as total costs are situated above the estimated cost frontier (Kumbhakar and Lovell, 2000).

Initially, many studies employed two-stage approach in which inefficiency estimates in the first step are regressed on a vector of potential correlates in the second step (see, for example, Berger and Mester, 1997; Chang et al, 1998). Two-stage approach was proved (by using Monte Carlo approach) to be inappropriate due to omitted variable bias by Wang and Schmidt (2002).

Alternatively, Kumbhakar et al (1991) proposed a single-step approach how to simultaneously estimate cost inefficiency and correlated variables in order to determine possible sources of its variation among different entities. Battese and Coelli (1995) extended this initially proposed model and also made it applicable for panel data.

The following equation can be inserted into a single-stage cost frontier model in order to estimate efficiency scores and correlates of inefficiency term u_{it} following Battese and Coelli (1995) approach:

$$u_{it} = (\sum_{j=1}^{J} \gamma' z_{jit} + e_{it}) \geq 0$$

Systematic component $\gamma' z_{jit}$ is associated with an arbitrary number of exogenous variables which indirectly influence composition of cost within associated production process when inputs are converted into outputs. In addition, as specified by Battese and Coelli (1995), $\gamma' z_{it}$ can be both positive or negative.

In this scenario non-negative inefficiency term which was defined in section 3.1 has no longer assumed to have a zero mean. Conditional mean (or mode) substitutes zero mean (or mode) and the updated distribution can be shown in a following way:

$$u_{it} \sim N^{+}(\gamma' z_{jit}, \sigma_u^2)$$

Finally, the remaining error term e_{it} in the above equation is distributed normally by assumption as follows:

$$e_{it} \sim N\ (0, \sigma_e^2)$$

The main weakness of the proposed Batesse and Coeli (1995) approach is that heteroscedasticity can be a serious concern in both inefficiency and random error terms. The most important part is that inefficiency scores can be biased upwards and the proposed model, unfortunately, cannot disentangle heteroscedasticity from estimated inefficiency. For example, larger banks might be more cost efficient, however, due to higher heterogeneity (related to, for example, different technology or economic behavior) which is incorporated in inefficiency term the results would become misleading.

4. Data

Data for this study are taken from the Bank Regulatory database provided by the Wharton Research Data Services (WRDS). Data are based on the Report of Condition and Income (Call Report) for all commercial banks in the U.S. that report to the Federal Reserve banks and the FDIC.

4.1. Data sample and descriptive statistics

Data covers the period from 2008 to 2013. Similar to the approach of Feng and Serletis (2009), only continuously operating banks are taken into account in order to analyze the performance of healthy and well-established institutions with no impact on entry or exit. A screening process is applied to eliminate banks with no loans, or no deposits, or no total operating expenses (consisting of total interest expense and total non-interest expense), or no personnel expenses (based on no data in salaries section). Branches and agencies of foreign banking organizations (FBOs) are entirely eliminated in the process because of no data on deposits and total operating expenses including interest expense, salaries and other non-interest expense.

According to Goulding and Nolle (2012), foreign-owned branches are usually prohibited from collecting retail deposits from individual clients (citizens or residents in the U.S.). Only few insured branches and agencies are exempted from this rule. Also, foreign-owned branches are required to hold certain amounts of high quality assets in the U.S. Finally, branches are more likely as a banking business model when foreign operations are smaller in size and do not have a retail orientation. These facts confirm that it is not meaningful to compare foreign-owned branches with domestic banks in the U.S. thus mainly foreign-owned banks having subsidiary status remain for comparison purposes with their domestic counterparts.

There are few acquisitions of domestic banks by foreign institutions during 2008-2013. If a bank is acquired during 2008-2010, I treat it as a foreign bank during the entire period of 2008-2013 while if a bank is acquired during 2011-2013, I treat it as a domestic bank in the same manner.

After the full screening process, I obtained a balanced panel of **6,080** banks out of which **38** banks are classified as FBOs. For descriptive statistics and classification of banks according to size, please refer to Table 1 and Table 2 in the appendix. For the full list of FBOs and top 30 largest banks in the sample sorted by average total assets during 2008-2013, refer to Table 3 and Table 4 in the appendix.

4.2. Elements of cost frontier

There are two main approaches how to define outputs and inputs in SFA. The first approach is called intermediation (or asset) approach introduced by Sealey and Lindley (1977) where loans, securities and other assets are treated as outputs while deposits are included in an input price estimation (i.e., borrowed funds are interest expense over deposits). Many researchers in the field employed this method when dealing with banking industry (for example, Mester, 1997; Berger and Mester, 2003; Greene, 2005a and 2005b; Feng and Serletis, 2009).

The second method is called value-added (also known as production or activity) approach defined by Berger et al (1987) where outputs are determined by the criteria of value added from most significant labor and capital expenditure. Loans and deposits are considered as outputs (or products) under this approach. Research under value-added approach in banking industry is also extensive (see, for example, Berger and Humphrey, 1991; Fries and Taci, 2005; Poghosyan and Kumbhakar, 2010).

Both approaches have their advantages and drawbacks. The value-added approach is considered the most accurate when dealing with technology changes and efficiency over time (Berger and Humphrey, 1992). The asset approach is intuitively correct as it separates output-oriented asset side from input-oriented liability side. The latter approach works best in cases banks purchase deposits from other banks or customers and subsequently turn them into loans. However, this is often not the case as usually banks also provide substantial services to their depositors (Berger and Humphrey, 1992). Thus there is no clear answer whether deposits should be treated as inputs or outputs.

I decided to test the value-added approach on U.S. banking industry. I identify two outputs which are total loans (both consumer and non-consumer loans) and total deposits, and two input prices which are capital and labor. The price of capital is equal to non-interest expense excluding labor expense divided by premises and other fixed assets (both tangible and intangible assets). I define labor input as labor expense over total assets. Total costs include both interest and non-interest expenses. The presented elements of cost frontier are summarized in the Table I below.

Table I. Input and output variables

Variable	**Symbol**	**Name**	**Description**
Dependent variable	TC	Total cost	Sum of interest and non-interest (labor and other non-interest) expenses
Output	Y_1	Loans	Sum of long-term and short-term loans
Output	Y_2	Deposits	Sum of deposits
Input price	X_1	Price of capital	Other non-interest expenses (excluding labor) over fixed assets (including intangible assets)
Input price	X_2	Price of labor	Labor expenses over total assets

A generic translog specification form with defined multiple outputs and inputs can be shown below:

$$\ln(\mathrm{TC}/\mathrm{X}_{\mathrm{Labor}}) = \beta_0 + \alpha_1 \ln \mathrm{Y}_{\mathrm{Loans}} + \alpha_2 \ln \mathrm{Y}_{\mathrm{Deposits}} +$$

$$\frac{1}{2}\sum_{k=1}^{2}\sum_{m=1}^{2} \alpha_{km} \ln \mathrm{Y}_{\mathrm{Loans}} \ln \mathrm{Y}_{\mathrm{Deposits}} +$$

$$\beta_1 \ln(\mathrm{X}_{\mathrm{Capital}}/\mathrm{X}_{\mathrm{Labor}}) + \frac{1}{2}\beta_1 \ln(\mathrm{X}_{\mathrm{Capital}}/\mathrm{X}_{\mathrm{Labor}})^2 +$$

$$\sum_{k=2}^{2}\sum_{l=1}^{1}\varphi_{kl}\ln Y_k \ln\left(X_{Capital}/X_{Labor}\right) + v_{it} + u_{it}$$

Notice that normalization of the dependent variable and input prices is applied in the formula - division by one of input prices (labor price in this case) imposes linear homogeneity in total costs and capital price. This is an ordinary practice in similar research (Kumbhakar and Lovell, 2000).

4.3. Bank-specific inefficiency correlates

Bank-specific inefficiency correlates are collected in order to reflect several bank characteristics such as leverage, risk, size, ownership, diversification, competitiveness and output quality indicators. These variables are also called exogenous variables because they tend to define an environment of banking business processes. They should influence efficiency of how banking inputs are converted into outputs. In order to select inefficiency correlates, I consider all relevant research discussed in literature review section. According to Maudos et al (2002), there is no theoretical model of what to include as potential correlates thus it is not appropriate to call them explanatory variables but rather inefficiency correlates. The full list of exogenous variables applied in this study including expected signs towards inefficiency based on previous research (for example, Berger and Mester, 1997; Fries and Taci, 2005, etc.) is provided in a Table II below.

Table II. Bank-specific inefficiency correlates

Name	**Exp. sign**	**Bank characteristic**	**Description**
Inverse leverage ratio (equity-to-assets)	(+)	Risk / leverage	This ratio captures risk preferences of individual banks over time

Loan loss provisions-to-loans	(+)	Output quality	This ratio reflects performance of lending services over time (what proportion of loans are considered as non-recoverable in terms of total loan portfolio)
Non-interest income-to-total assets	(-)	Diversification	This ratio reflects a degree of diversification of banking services
Net interest margin	(+/-)	Competitiveness/ market power	Measurement of competition level over time calculated as net interest income to total assets. Also, this ratio reflects quality of loan portfolio.
FBO dummy	(+/-)	Ownership	Ownership dummy of an individual bank (1 if foreign owned, 0 if domestically-owned)
Large dummy	(-/+)	Size	Size dummy of an individual bank (1 if average total assets over 6 years are greater than $1 billion, 0 if otherwise)
Medium dummy	(-/+)	Size	Size dummy of an individual bank (1 if average total assets over 6 years are between $0.3 billion and $1 billion, 0 if otherwise)

The ownership dummy is of particular importance of this study. I am going to examine how an impact of a foreign-owned bank in the sample and sub-samples of large and medium banks affects the business environment of banking industry in the U.S. I will also be able to compare ownership significance in relation to other exogenous variables.

5. Empirical results

The results estimated using Battese and Coelli (1995) specification are provided in the appendix. Cobb-Douglas specification is illustrated in Table 5 while translog specification is shown in Table 6. All continuous explanatory variables and the dependent variable are in natural logarithms.

5.1. Results from Cobb-Douglas cost function

The analysis of obtained results starts from Cobb-Douglas specification which is easier to estimate and interpret, and there are less parameters than in a more flexible translog form. To be more specific, Cobb-Douglas function is nested in translog specification as shown in a methodology part.

To begin with, cost function model with two outputs (loans and deposits) and normalized input price (capital over labor) show positive and highly statistically significant results. Loan coefficient is more than three times smaller than deposit coefficient. This indicates that 1% increase in total loans would lead to about three times less percentage increase in dependent variable (total costs normalized by labor price) than in case of identical 1% increase in total deposits. Total costs are composed of a significant part of interest expense which come from deposits thus such results are reasonable and expected.

Furthermore, in order to check robustness of the first equation and to account for the fact that this base equation violated curvature requirements due to multiple outputs, I also estimate single-output Cobb-Douglas equations for loans and for deposits separately. I find positive and significant results for both loans and deposits when each of them are selected as a single output.

Due to the fact that both single-output coefficients and sum of the output coefficients in the base case scenario are approximately equal to 1, such outcome can be interpreted as nearly constant economies of scale for an average bank in the sample.

I also test two Cobb-Douglas specifications for sub-samples of large banks (with average total assets exceeding USD 1 billion during the sample period) and medium banks (with average total assets between USD 0.3 billion and USD 1 billion during the sample period). The results are again highly significant with outputs showing slight diseconomies of scale, especially in case of medium banks.

Gamma parameter shows which part of variance in the composed error term is coming from inefficiency component. Gamma parameter values in estimated results are in the range 0.93-0.99 and these numbers indicate that inefficiency term quite strongly dominates random error. However, functional form misspecification might be a possible issue when random error term is almost non-existent, thus it is necessary to test more flexible model such as general translog function.

5.2. Results from translog cost function

The analysis of translog specification (quadratic in logs) is, unfortunately, more difficult to interpret as it involves many parameters (in this case it involves 12 parameters while Cobb-Douglas involves 6 parameters, excluding inefficiency correlates) with second order approximations. I use likelihood ratio test for stochastic frontier models in order to compare if translog specifications are different from nested Cobb-Douglas functional forms. Null hypothesis that translog model can be reduced to Cobb-Douglas is rejected in all cases. For the results, please refer to Table 7.

The results of translog cost function are in majority significant at 1% level. Again, the effect on total costs of 1% increase in deposits is multiple times larger than 1% increase in loans which is reasonable due to the structure of total costs. The cross-output term between loans and deposits is negative and statistically significant. The following result might show possible economies of scope between lending and borrowing activities for an average bank in the sample.

In order to compare the stochastic frontiers of all banks in the sample with large banks (above $1 billion) and medium banks (between $0.3 billion and $1 billion) three additional translog cost functions are analyzed (refer to Table 6 in the appendix). Significance of individual parameters in each of the models slightly decline which is in line with expectations taking into account that translog specification is more prone to multicollinearity issues for smaller samples. Additionally, some interesting findings regarding inefficiency correlates are interpreted in the sub-section 5.3.

Gamma parameter values in estimated results are in the range 0.88-0.99 and these numbers again indicate that inefficiency term is more important than any other stochastic variation in these models. Total sum of variance of both error components based on sigma-squared estimate is high for large banks which might be due to rather small sample of 505 banks.

Overall, sigma-squared decreases for all translog models comparing to Cobb-Douglas method. Based on other authors work (see, for example, Feng and Serletis, 2009), functional form of stochastic frontier actually differs for different group of banks. Thus such results might reflect that functional form is better represented in translog specification as opposed to Cobb-Douglas functional form.

5.3. Results from the analysis of inefficiency correlates

In this sub-section I will turn to analysis of inefficiency correlates with a particular focus on investigation what drives inefficiency differences among individual banks.

Inverse leverage ratio (equity-to-assets) is significantly positively correlated with cost inefficiency and all results consistently show this relationship. Such results are intuitive because banks must sacrifice a portion of cost efficiency in order to satisfy regulatory capital requirements and individual risk management procedures. For example, more assets will be permanently shifted to equity instead of being provided as loans in case of more risk-averse banks.

Loan loss provisions over loans ratio (similarly to inverse leverage ratio) is significantly positively correlated with cost inefficiency. Such results are also expected as estimated losses due to inability to recover money in lending business must have a negative impact on cost efficiency. Worse performing banks might have taken more risk in their lending activities or they might be underperforming in allocation of loans. In addition, a negative effect of this ratio on cost efficiency was more than multiple times higher for larger banks comparing to medium banks in the U.S. Thus large banks suffered relatively more in terms of cost efficiency in case of higher loan losses.

Non-interest income-to-total assets which is a proxy of diversification of banking activities shows mixed results. This ratio is significantly negatively associated with cost inefficiency for medium banks which means medium banks benefit from higher business diversification through wider variety of banking services. On the other hand, the analysis shows completely opposite and significant results for large banks. It would be important to also test the relation of non-interest income-to-assets in terms of profit efficiency in order to understand if some less cost efficient banking services are also less (or not) profit efficient for large banks in line with Berger and Mester (1997).

Net interest margin also shows different results taking into account sub-sample of large banks in contrast to medium banks and all the rest banks in the sample. This ratio is positively associated with cost inefficiency for large banks, although the results are significant only at 10% level in translog specification. This result can be explained that banks with higher market power should be able to decrease their costs comparing to banks which earn only a marginal net interest and experience a gain in terms of cost efficiency via increased scale of services. In line, this ratio shows opposite results for medium and smaller banks (as reflected in full sample results) showing that for banks with less market power a drop in net interest margin will be associated with declining cost efficiency.

I also use dummies in order to track how size of banks in terms of average total assets (during the 6 years' sample period) affect cost efficiency. I propose dummies for large banks (above USD 1 billion) and medium banks (between USD 0.3 billion and USD 1 billion). The results actually show that both large banks and medium banks are positively associated with cost inefficiency - the opposite of that might be expected. Large banks are in fact the most cost inefficient in the sample as shown in the translog specification (based on large dummy coefficient in a sub-sample including both large and medium banks but excluding small banks). I find that my results regarding size dummies are in line with Feng and Serletis (2009) and Kaparakis (1994) who use a similar sample size of banks in the U.S. during different decades and also find that largest banks are the least cost efficient.

5.4. Cost efficiency and foreign ownership

Foreign ownership analysis is a key focus of this study. Considering all except one scenarios of Cobb-Douglas specification (see Table 5) I find that foreign ownership is significantly positively correlated with inefficiency term. The interpretation of such results is that foreign banks on average operate less cost efficiently than their domestic counterparts. This finding would actually reflect general consensus that domestic banks in developed economies such as the U.S. are more cost efficient (Berger (2007), Lensink et al (2008), Lensink and Meesters (2014).

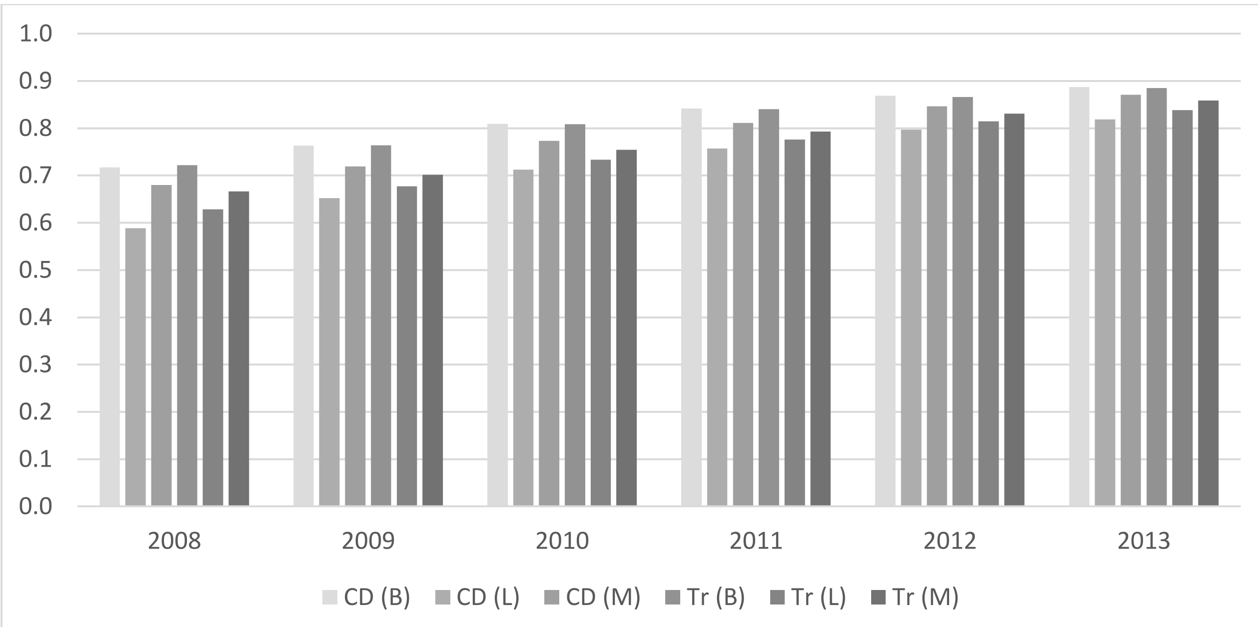
In opposite, the results from translog specification (see Table 6) indicate a completely different picture. The base specification shows that foreign ownership negatively correlates with cost efficiency, however, this result is not significant. As the majority of foreign banks in the sample exceed $0.3 billion of average total assets during 2008-2013 (see Table 2 in the appendix), it is important to use additional sub-samples reflecting only large or medium banks and test foreign ownership among banks of similar size. Translog specification of a sub-sample of large banks exceeding USD 1 billion shows that foreign banks are less cost efficient, but only at 10% significance level. On the other hand, foreign banks in medium banks' sub-sample are more cost efficient, but without any significance. Finally, no significance is found in case of both large and medium banks in a sub-sample.

Therefore, the overall outcome of whether foreign banks and domestic banks significantly differ in terms of cost efficiency is inconclusive. Additional analysis and methods are needed in order to better understand this relationship. It is also worth noting that cost efficiency is only one side of a full picture and it would be important to also study profit efficiency in order to better grasp outstanding differences between domestic and foreign banks.

5.5. Bank efficiency scores

Chart I below reflects an increasing trend of average efficiencies on a yearly basis since the financial crisis of 2008-2009. Full sample base scenario, sub-samples of large banks and medium banks are included for Cobb-Douglas and translog specifications in the chart. Mean efficiencies for the overall period are provided in Table 5 and Table 6 in the appendix.

Chart I: Mean efficiency by each year



Bank efficiency scores reflect tendency that larger banks are the least cost efficient while medium banks are less cost efficient comparing to small banks. However, this result might show that significant part of hererogeneity due to different technology is included in inefficiency component for larger banks. For example, Orea and Kumbhakar (2004) review several papers using SFA in banking research and show that technological differences usually occur in inefficiency term resulting in an upward-bias. Unfortunately, it is not methodologically possible to disentangle heterogeneity component under Battese and Coelli (1995) specification and then compare the residual inefficiency for banks of different sizes.

Other possible scenario is that larger banks are less cost efficient but actually more profit efficient. Larger banks might use their market power in order to participate in more profitable (but less cost efficient) activities.

5.6. Limitations and further research

The main limitation of this study, in my opinion, is that inefficiency term cannot be analyzed further according to the used methodology in order to make sure that it does not contain heterogeneity component. The proposition for further research would be to consider more advanced techniques currently on the trend in the field of stochastic frontier analysis including latent class models, local maximum likelihood method and some very recent advances including Bayesian approach (see the discussion in Kumbhakar and Tsionas, 2011).

The other important drawback in parametric SFA method which constantly receives attention is flexibility issue (Feng and Serletis, 2009; Kumbhakar and Tsionas, 2011). One of possible solutions would be to test more flexible functional forms, for example, translog function with augmented Fourier trigonometric terms (for details, see Berger and Mester, 1997; 2003). Based on the review of recent advances by Parmeter and Kumbhakar (2014), the most promising area, according to the authors, is semi-parametric and nonparametric estimation of stochastic frontier models. Such methods can have a better balance in solving the most prominent issues in SFA and DEA approaches which are imposition of a functional form and low tolerance of random noise in an analyzed data.

Finally, in order to have a well-rounded picture of banking efficiency in the U.S., cost efficiency analysis should be supplemented by insights from profit efficiency studies. Also, banks could be compared how they perform in different states and what are state-specific drivers influencing changes in cost and profit efficiency.

6. Conclusion

The banking sector in the U.S. has undergone changing regulatory environment, high level of consolidation and technological transformation during the last three decades. Foreign banks in the U.S. were affected by all these processes as well; also, there might be further attempts to increase regulation and negatively affect efficiency of this banking segment. Thus there is a need to investigate the current state of performance of foreign banks and benchmark them to domestic banks in terms of ability to reduce costs in highly regulatory environment after the financial crisis.

This study, therefore, provides some insights about foreign ownership and cost efficiency of banks in the U.S. I employ stochastic frontier analysis which allows to estimate cost efficiency functions, calculate average cost efficiency scores and find several significant bank-specific factors which determine cost inefficiency. Battese and Coelli (1995) one-step parameterization is applied in order to obtain more consistent results of inefficiency correlates. The effects of selected inefficiency correlates are investigated simultaneously in order to limit potential omitted variable bias which is evident in sub-optimal two-step analysis.

The average cost efficiency (the results are provided in Table 5 and Table 6 in the appendix) in the full sample is estimated to be 81%. When subsamples of large banks and medium banks are analyzed, average cost efficiency decreases to the range of 72% to 74% for large banks and 77% to 78% for medium banks. The increasing trend of average cost efficiency since the financial crisis is observed over the investigated period of 2008-2013, which is detected for both full sample and sub-samples based on bank size.

Empirical results show and allow to interpret significant relations of cost inefficiency with exogenous variables leverage (risk), size, diversification of banking services, output quality and performance (market power) indicators. Based on results, adding more equity in order to decrease riskiness makes banks less cost efficient. Larger banks were found to be less cost efficient but this finding must be also investigated in terms of profit efficiency in order to draw further conclusions. As expected, output quality in terms of proportion of loan loss provisions-to-total assets is positively associated with cost inefficiency. Diversification (non-interest income-to-total assets) and performance (net interest margin, also might refer to a proxy of market power) indicators show mixed results for banks of different sizes.

Finally, foreign ownership is revealed to have no significant influence on cost efficiency which could be observable over all different models under investigation. Other inefficiency correlates mentioned above are better at explaining what drives cost efficiency for different banks in the U.S. I also account for the fact that this study is limited due to the model constraints and the sole focus on cost efficiency, therefore, I encourage further research and advise to extend the analysis of foreign ownership in more advanced settings which could hopefully reveal further important insights regarding this topic.

References

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

Battese, G., and Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. Journal of Productivity Analysis, 3, 153–169.

Battese, G.E., and Coelli, T. J. (1995). A Model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics, 20(2), 325-332.

Belotti, F., Daidone, S., Ilardi, G., and Atella, V. (2013). Stochastic frontier analysis using Stata, The Stata Journal, 13(4), 719-758.

Belotti, F., and Ilardi, G. (2015). Consistent inference in fixed-effects stochastic frontier models. Temi di discussione (Economic working papers), Bank of Italy (forthcoming).

Berger, A. N. (2007). International comparisons of banking efficiency. Financial Markets, Institutions & Instruments, 16(3), 119-144.

Berger, A. N., Hanweck, G. A., and Humphrey, D. B. (1987). Competitive viability in banking: scale, scope, and product mix economies. Journal of Monetary Economics, 20(3), 501-520.

Berger, A. N., and Humphrey, D. B. (1991). The dominance of inefficiencies over scale and product mix economies in banking. Journal of Monetary Economics, 28(1), 117–148.

Berger, A. N., and Humphrey, D. B. (1992). Measurement and efficiency issues in commercial banking. Output Measurement in the Service Sectors. The National Bureau of Economic Research in NBER Book Series Studies in Income and Wealth, 56, 245-300. University of Chicago Press, Chicago.

Berger, A. N., Humphrey D. B. (1997). Efficiency of financial institutions: international survey and directions for future research. European Journal of Operational Research, 98(2), 175–212.

Berger A. N., and Mester L. J. (1997). Inside the black box: what explains differences in the efficiencies of financial institutions? Journal of Banking and Finance, 21(7), 895–947.

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

Berger, A. N., Young, R., Genay, H., and Udell, G. F. (2000). Globalization of Financial Institutions: Evidence from Cross-Border Banking Performance. Brookings-Wharton Papers on Financial Services, 2000-04.

Berlin, M. (2015). New rules for foreign banks: what's at stake? Business Review, Q1, 1-10.

Bonin, J., Hasan, I., and Wachtel, P. (2005). Bank privatization and performance: evidence from transition countries. Journal of Banking and Finance 29, 31–53.

Cerutti, E., Dell'Ariccia, G., and Martinez Peria, M. S. (2007). How banks go abroad: branches or subsidiaries? Journal of Banking and Finance, 31, 1,669-1,692.

Chang, C. E., Hasan, I., and Hunter, W. C. (1998). Efficiency of multinational banks: an empirical investigation. Applied Financial Economics, 8(6), 689-696.

Chen, Y., Wang, H., and Schmidt, P. (2014). Consistent estimation of the fixed effects stochastic frontier model. Journal of Econometrics, 181, 65–76.

Coelli, T. J., and Henningsen, A. (2013). Frontier: stochastic frontier analysis. R package version 1.0. http://CRAN.R-Project.org/package=frontier.

Cornwell, C., Schmidt, P., and Sickles, R. C. (1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. Journal of Econometrics, 46(1-2), 185–200.

DeYoung, R., and None, D. (1996). Foreign-owned banks in the U. S.: earning market share or buying it? Journal of Money, Credit, and Banking, 28(4), 622-636.

Federal Reserve Press Release, February 18, 2014. Retrieved from: https://www.federalreserve.gov/newsevents/press/bcreg/20140218a.htm

Feng, G., and Serletis, A. (2009). Efficiency and productivity of the US banking industry, 1998-2005: evidence from the Fourier cost function satisfying global regularity conditions. Journal of Applied Econometrics, 24, 105-138.

Fries, S., and Taci, A. (2005). Cost efficiency of banks in transition: evidence from 289 banks in 15 post-communist countries. Journal of Banking & Finance, 29(1).

Goulding, W., and Nolle, D. E. (2012). Foreign banks in the U.S.: a primer. Board of Governors of the Federal Reserve System. International Finance Discussion Papers, 1064.

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

Greene, W. (2005b). Fixed and random effects in stochastic frontier models. Journal of Productivity Analysis, 23, 7–32.

Hasenkamp, G. (1976). Specification and estimation of multiple-output production functions. New-York: Springer-Verlag.

Heckman, J. (1979). Sample selection bias as a specification error. Econometrica, 47, 153–161.

Huang, C.J., and Liu, J.-T. (1994). Estimation of a non-neutral stochastic frontier production function. Journal of Productivity Analysis, 5, 171-180.

Jondrow, J., Lovell, C., Materov, I., & Schmidt, P. (1982). On the estimation of technical efficiency in the stochastic production function model. Journal of Econometrics, 19, 233–238.

Kaparakis, E., Miller, S., and Noulas, A. (1994). Short-run cost inefficiency of commercial banks: a flexible stochastic frontier approach. Journal of Money, Credit, and Banking 26(4), 875–893.

Kumbhakar, S. C. (1990). Production frontiers, panel data and time-varying technical inefficiency. Journal of Econometrics, 46, 201–212.

Kumbhakar, S. C., Ghosh, S., and McGuckin, J. T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms. Journal of Business and Economic Statistics, 9(3), 279-86.

Kumbhakar, S. C. and Lovell K. C. A. (2000). Stochastic frontier analysis. Cambridge University Press, Cambridge.

Kumbhakar, S. C., and Tsionas, E. G. (2011). Some recent developments in efficiency measurement in stochastic frontier models. Journal of Probability and Statistics, 2011.

Lancaster, T. (2000). The incidental parameter problem since 1948. Journal of Econometrics, 95(2), 391–413.

Lensink, R., and Meesters, A. (2014). Institutions and bank performance: a stochastic frontier analysis. Oxford Bulletin of Economics and Statistics, 76, 67–92.

Lensink, R., Meesters, A., and Naaborg, I. (2008). Bank efficiency and foreign ownership: do good institutions matter? Journal of Banking & Finance, 32(5), 834-844.

Maudos, J., Pastor, J. M., Perez, F. and Quesada, J. (2002). Cost and profit efficiency in European banks. Journal of International Financial Markets, Institutions and Money, 12, 33-58.

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

Mester, L. J. (1997). Measuring efficiency at US banks: accounting for heterogeneity is important. European Journal of Operational Research, 98(2), 230–242.

Montgomery, L. (2003). Recent developments affecting depository institutions. FDIC Banking Review 15(2), 54–60.

Neyman, J. and Scott, E. L. (1948). Consistent estimates based on partially consistent observations. Econometrica. Journal of the Econometric Society, 16, 1–32.

Orea, L., and Kumbhakar, S. C. (2004). Efficiency measurement using a latent class stochastic frontier model. Empirical Economics, 29, 169–183.

Parmeter, F., and Kumbhakar, S. C. (2014). Efficiency analysis: a primer on recent advances. University of Miami, State University of New York at Binghamton.

Poghosyan, T., and Borovicka, J. (2007) Banking efficiency in emerging economies: the impact of ownership endogeneity and EU accession. CERGE-EI discussion paper no. 189.

Poghosyan, T., and Kumbhakar, S. C. (2010). Heterogeneity of technological regimes and banking efficiency in former socialist economies. Journal of Productivity Analysis, 33, 19-31.

Schmidt, P. and Sickles, R. (1984). Production Frontiers with Panel Data. Journal of Business and Economic Statistics, 2(4), 367-374.

Sealey C., and Lindley J. (1977). Inputs, outputs, and a theory of production and cost at depository financial institutions. Journal of Finance, 32(4), 1251–1266.

Wang, H., and Ho, C. (2010). Estimating fixed-effect panel stochastic frontier models by model transformation. Journal of Econometrics, 157, 286–96.

Wang, H., and Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. Journal of Productivity Analysis, 18, 129–144.

Appendix

Table 1: Descriptive statistics

Variable	Units	Mean	Std. Dev.	Min	Max
Total costs (TC)	USD million	79	1,333	0	64,635
Loans (Y_1)	USD million	1,077	17,605	0	788,796
Deposits (Y_2)	USD million	1,468	26,250	0	1,326,036
Price of capital (X_1)	%	14.81	792.48	0.00	89,702.00
Price of labor (X_2)	%	0.02	0.01	0.00	0.40
Inverse leverage ratio (equity-to-assets) (Z_1)	%	0.11	0.04	-0.04	0.96
Loan loss provisions-to-loans ratio (Z_2)	%	0.01	0.01	-0.23	0.43
Non-interest income-to-total assets (Z_3)	%	0.01	0.02	-0.12	0.79
Net interest margin (Z_4)	%	0.03	0.01	-0.01	0.29
FBO dummy (Z_5)	0 or 1	0.01	0.08	0	1
Large dummy (Z_6)	0 or 1	0.08	0.28	0	1
Medium dummy (Z_7)	0 or 1	0.20	0.40	0	1
Total assets	USD million	2,123	38,357	3	1,945,467

Table 2: banks by size class based on total assets (6 years average during 2008-2013)

	Description	Bank size (B$)	# of banks in total	# of FBOs
1	Very small	<0.1B	2,017	0
2	Small	<0.3B	2,328	2
3	Medium	<1B	1,230	8
4	Large	<10B	428	15
5	Very large	>10B	77	13
	Total		**6,080**	**38**

Table 3: list of FBOs sorted by total assets (6 years average during 2008-2013)

	Name	City	State	Previous name	TA (M$)
1	HSBC BK USA NA	McLean	VA		183,743
2	T D BK NA	Wilmington	DE		170,157
3	RBS CITIZENS NA	Providence	RI		110,161
4	UNION BK NA	San Francisco	CA		87,364
5	BMO HARRIS BK NA	Chicago	IL	HARRIS NA	74,187
6	COMPASS BK	Birmingham	AL		65,679
7	BANK OF THE WEST	San Francisco	CA		62,794
8	DEUTSCHE BK TC AMERICAS	New York	NY		50,903
9	UBS BK USA	Salt Lake City	UT		36,502
10	CITIZENS BK OF PA	Philadelphia	PA		32,498
11	BARCLAYS BK DE	Wilmington	DE		15,446
12	FIRST HAWAIIAN BK	Honolulu	HI		15,219
13	RABOBANK NA	Roseville	CA		11,411
14	ISRAEL DISCOUNT BK OF NY	New York	NY		9,556
15	GREAT WESTERN BK	Sioux Falls	SD		7,448
16	MERCANTIL COMMERCEBANK NA	Coral Gables	FL		6,462
17	BANCO SANTANDER INTL	Miami	FL		6,461
18	HSBC PRIVATE BK INTL	Miami	FL		5,574
19	BANK LEUMI USA	New York	NY		5,386
20	CITY NB OF FL	Miami	FL		4,387
21	SABADELL UNITED BK NA	Miami	FL	MELLON UNITED NB	2,952
22	TOTALBANK	Miami	FL		2,209
23	MANUFACTURERS BK	Los Angeles	CA		2,074
24	INTER NB	McAllen	TX		2,032
25	CTBC BK CORP USA	Los Angeles	CA	CHINATRUST BK USA	1,837
26	FAR EAST NB	Los Angeles	CA	FAR E NB	1,634
27	BANCO ITAU INTL	Miami	FL	BANCO ITAU EUROPA INTL	1,581
28	WOORI AMER BK	New York	NY		1,045
29	SHINHAN BK AMER	New York	NY		966
30	STATE BK OF INDIA CA	Los Angeles	CA		786
31	INDUSTRIAL & CMRL BK	New York	NY	BANK OF EAST ASIA USA NA	767
32	ESPIRITO SANTO BK	Miami	FL		607
33	EVERTRUST BK	Pasadena	CA		530
34	DEUTSCHE BK TC DE	Wilmington	DE		474
35	FIRST CMRL BK USA	Alhambra	CA		462
36	PACIFIC NB	Miami	FL		378
37	DESJARDINS BK NA	Hallandale	FL		181
38	NATBANK NA	Hollywood	FL		117

Table 4: top 30 largest banks sorted by total assets (6 years average during 2008-2013)

	Name	City	State	FBO	TA (M$)
1	JPMORGAN CHASE BK NA	Columbus	OH		1,776,578
2	BANK OF AMER NA	Charlotte	NC		1,464,121
3	CITIBANK NA	Sioux Falls	SD		1,248,583
4	WELLS FARGO BK NA	Sioux Falls	SD		1,008,538
5	JP MORGAN INTL FNC	Newark	DE		462,251
6	CITIBANK OVERSEAS INV CORP	New Castle	DE		403,921
7	U S BK NA	Cincinnati	OH		312,741
8	PNC BK NA	Wilmington	DE		254,344
9	BANK OF NY MELLON	New York	NY		229,418
10	STATE STREET B&TC	Boston	MA		191,749
11	HSBC BK USA NA	McLean	VA	Yes	183,743
12	SUNTRUST BK	Atlanta	GA		170,597
13	T D BK NA	Wilmington	DE	Yes	170,157
14	CAPITAL ONE NA	McLean	VA		165,388
15	BRANCH BKG&TC	Winston Salem	NC		164,003
16	FIA CARD SVC NA	Wilmington	DE		162,365
17	REGIONS BK	Birmingham	AL		128,144
18	CHASE BK USA NA	Wilmington	DE		115,066
19	GOLDMAN SACHS BK USA	New York	NY		111,925
20	RBS CITIZENS NA	Providence	RI	Yes	110,161
21	FIFTH THIRD BK	Cincinnati	OH		108,890
22	KEYBANK NA	Cleveland	OH		90,720
23	UNION BK NA	San Francisco	CA	Yes	87,364
24	NORTHERN TC	Chicago	IL		84,874
25	BMO HARRIS BK NA	Chicago	IL	Yes	74,187
26	MANUFACTURERS & TRADERS TC	Buffalo	NY		73,843
27	ALLY BK	Midvale	UT		72,897
28	COMPASS BK	Birmingham	AL	Yes	65,679
29	CAPITAL ONE BK USA NA	Glen Allen	VA		64,113
30	DISCOVER BK	Greenwood	DE		63,683

Table 5: Results from Cobb-Douglas functional form

Dependent variable: Ln_(total costs/labor)	Base scenario	Single output (loans)	Single output (deposits)	Large banks (>1B$)	Medium banks [.3B$, 1B$]
Intercept	.8164***	1.4069***	.6582***	.1869**	2.0330***
	.0157	.0198	.0156	.0755	.0924
Ln_loans	.2354***	.9556***		.3815***	.3290***
	.0039	.0015		.0253	.0128
Ln_deposits	.7637***		1.0090***	.6436***	.5692***
	.0045		.0011	.0253	.0141
Ln_(capital/labor)	.0509***	.0796***	.0379***	.1251***	.0756***
	.0014	.0017	.0013	.0072	.0032
Z_0 - intercept	-22.971***	1.7770***	-28.288***	-23.604***	-.1641
	3.3349	.0691	3.5341	7.4922	.1801
Z_1 - inv. leverage ratio (equity-to-assets)	27.389***	1.1778***	32.732***	40.884***	2.8382***
	3.2870	.2271	3.6883	11.502	.4675
Z_2 - loan loss provision/loans	74.472***	15.454***	84.821***	33.901***	21.654***
	8.0588	1.2458	7.7163	12.635	2.5438
Z_3 - non-interest income/total assets	5.7696***	4.3531***	2.1462	29.643***	-6.6547***
	1.2551	.3642	2.3796	6.4672	1.2318
Z_4 - net interest margin	-22.000***	-207.74***	15.397***	53.285***	-32.371***
	6.1210	11.095	5.1773	18.200	4.9144
Z_5 - FBO dummy	**1.2287*** **	**-.5179*** **	**1.8403*** **	**2.7322*** **	**.4467*** **
	.2644	**.0857**	**.5389**	**1.0325**	**.1213**
Z_6 - Large dummy	10.467***	1.9193***	11.568***		
	1.4910	.1275	1.4863		
Z_7 - Medium dummy	4.7284***	.2200***	5.9688***		
	.7485	.0648	.7837		
Sigma-squared $\sigma^2 = \sigma_u^2 + \sigma_v^2$	4.1399***	1.3781***	5.1572***	6.7077***	.3615***
	.5751	.0925	.6452	2.1325	.0528
Gamma $\gamma = \sigma_u^2/(\sigma_v^2 + \sigma_u^2)$	.9934***	.9653***	.9946***	.9950***	.9341***
	.0009	.0026	.0007	.0017	.0084
Log-likelihood value	-2,826	-10,382	-4,314	-1,358	-639
Mean efficiency	.8146	.7858	.8041	.7210	.7836
N of observations	36,480	36,480	36,480	3,030	7,380
N of banks	6,080	6,080	6,080	505	1,230

Notes: *, **, *** - coefficients are significant at 10%, 5% and 1%, respectively (two-sided significance level, t-statistics). Standard errors of respective coefficients are shown below respective coefficients.

Table 6: Results from translog functional form

Dependent variable: Ln_(total costs/labor)	Base scenario	Large banks (>1B$)	Medium banks [.3B$, 1B$]	Large and medium banks
Intercept	1.8444*** .0930	5.2210*** .5423	-.2355 1.1591	5.2427*** .2461
Ln_loans	.1424*** .0437	-.5012** .2104	-.5190** .2200	-.3076*** .0998
Ln_deposits	.7852*** .0472	.9878*** .2082	1.8360*** .2640	.7668*** .1025
Ln_(capital/labor)	-.2529*** .0091	-.3202*** .0523	-.2341*** .0870	-.2346*** .0253
½ $(Ln\_loans)^2$	.1278*** .0031	.3046*** .0223	.4073*** .0260	.3165*** .0124
½ $(Ln\_deposits)^2$	.1109*** .0080	.1839*** .0143	.2033*** .0196	.2085*** .0085
½ $(Ln\_capital/labor)^2$	.0144*** .0009	.0188*** .0020	.0062** .0028	.0147*** .0017
Ln_loans x Ln_deposits	-.1199*** .0051	-.2295*** .0123	-.3237*** .0176	-.2464*** .0075
Ln_loans x Ln_(capital/labor)	.0246*** .0021	-.0186** .0073	-.0134 .0086	-.0344*** .0037
Ln_deposits x Ln_(capital/labor)	-.0032 .0022	.0410*** .0064	.0349*** .0095	.0529*** .0033
Z_0 - intercept	-1.9510*** .2689	-18.320*** 7.106	.4401*** .0701	-13.258*** 3.5433
Z_1 - inv. leverage ratio (equity-to-assets)	4.7402*** .4298	30.206*** 10.095	.5111** .2176	14.587*** 3.3169
Z_2 - loan loss provision/loans	23.553*** 1.5242	46.039*** 16.075	13.247*** 1.0527	47.921*** 9.2846
Z_3 - non-interest income/total assets	.3435 .2445	18.790*** 5.842	-6.7428*** 1.0262	9.4335*** 2.7888
Z_4 - net interest margin	-50.607*** 4.4424	22.133* 12.425	-19.842*** 2.1606	-9.5297* 5.5506
Z_5 - FBO dummy	**-.1254** **.1047**	**1.6620*** **.9006**	**-.1339** **.0936**	**.3823** **.3773**
Z_6 - Large dummy	1.6874*** .1503			4.9478*** 1.2720
Z_7 - Medium dummy	.7212*** .0719			
Sigma-squared $\sigma^2 = \sigma_u^2 + \sigma_v^2$	.7007*** .0614	4.8230*** 1.8352	.1843*** .0158	2.5587*** .6453
Gamma $\gamma = \sigma_u^2/(\sigma_v^2 + \sigma_u^2)$	.9602*** .0034	.9925*** .0029	.8787*** .0112	.9868*** .0033
Log-likelihood value	-2,072	-1,133	-489	-1,853
Mean efficiency	.8143	.7447	.7676	.7981
N of observations	36,480	3,030	7,380	10,410
N of banks	6080	505	1,230	1,735

Table 7: Log-likelihood tests

Test (H_0 vs H_1)	Log-likelihood values	Degrees of freedom	Chi-sq.	Probability (>chi-sq.)
OLS (no inefficiency) vs	-9,465	5		
Cobb-Douglas (base sc.)	-2,826	14	13,278	.000***
OLS (no inefficiency) vs	-5,987	11		
translog (base sc.)	-2,072	20	7,830	.000***
Cobb-Douglas (loans) vs	-10,382	13		
Cobb-Douglas (base sc.)	-2,826	14	15,111	.000***
Cobb-Douglas (deposits) vs	-4,314	13		
Cobb-Douglas (base sc.)	-2,826	14	2,975	.000***
Cobb-Douglas (base sc.) vs	-2,826	14		
translog (base sc.)	-2,072	20	1,508	.000***
Cobb-Douglas (large) vs	-1,358	12		
translog (large)	-1,133	18	449	.000***
Cobb-Douglas (medium) vs	-639	12		
translog (medium)	-489	18	300	.000***

*** Asterisks on the value of probability (>chi-sq.) indicate that it exceeds the 99.9th percentile for the corresponding chi-squared distribution and so the null hypothesis that the unrestricted model can be reduced to restricted model is rejected in all cases.