ISSN 1518-3548

BANCO CENTRAL DO BRASIL

Working Paper Series 275

A geographically weighted approach in measuring efficiency in panel data: the case of US saving banks

Benjamin M. Tabak, Rogério B. Miranda and Dimas M. Fazio

April, 2012

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	n. 275	Apr.	2012	p. 1-33

Working Paper Series

Edited by Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Benjamin Miranda Tabak – E-mail: benjamin.tabak@bcb.gov.br
Editorial Assistant: Jane Sofia Moita – E-mail: jane.sofia@bcb.gov.br
Head of Research Department: Adriana Soares Sales – E-mail: adriana.sales@bcb.gov.br

The Banco Central do Brasil Working Papers are all evaluated in double blind referee process.

Authorized by Carlos Hamilton Vasconcelos Araújo, Deputy Governor for Economic Policy.

General Control of Publications

Banco Central do Brasil

Secre/Comun/Cogiv

SBS – Quadra 3 – Bloco B – Edifício-Sede – 1º andar

Caixa Postal 8.670

70074-900 Brasília – DF – Brazil

Phones: +55 (61) 3414-3710 and 3414-3565

Fax: +55 (61) 3414-3626

E-mail: editor@bcb.gov.br

Consumer Complaints and Public Enquiries Center

Banco Central do Brasil

Secre/Comun/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Fax: +55 (61) 3414-2553

Internet: <http://www.bcb.gov.br/?english>

A geographically weighted approach in measuring efficiency in panel data: the case of US saving banks

Benjamin M. Tabak* Rogério B. Miranda† Dimas M. Fazio‡

Abstract

The objective of this article is to discuss a new approach to control for the environment when one estimates efficiency by the stochastic frontier model. By introducing geographical weights and estimating local frontiers for each US saving bank for 2001-09, we find that bank technical performance is higher for most banks in comparison to a fixed-effects approach. This result highlights the importance of explicitly considering local environment and constraints while analyzing banks' behavior. All in all, this model has been proved very promising and viable for future empirical studies.

Key Words: Stochastic Frontier, Environmental Variables, Bank Performance, Geographical weights

JEL Classification: G21, G28

*DEPEP, Banco Central do Brasil

†Departament of Economics, Universidade Católica de Brasília and Instituto de Pesquisa Econômica Aplicada - IPEA

‡Departament of Economics, Universidade de São Paulo

1 Introduction

The estimation of bank efficiency has become a recurrent subject of analysis in the literature (Berger et al., 2009; Tecles and Tabak, 2010; Fang et al., 2011). In fact, one can clearly identify this awareness in the development and adjustment of several methods for the estimation of this variable in a particular banking industry[1]. A further motivation for the study of this matter is the recent financial crisis, which had clear impact on worldwide banking performance and stability. In this paper, we propose to estimate technical efficiency for US Saving Banks over the period of 2001-2009. Since the US financial system has the peculiarity of being connected with the whole world, its well-functioning is of uttermost importance. We basically use a modification of the stochastic frontier analysis (SFA) in order to calculate these efficiencies. The SFA is a parametric approach that estimates a frontier for a set of banking systems and compare each bank in the sample to it.

One interesting conclusion of the bank efficiency literature is that environmental conditions play a significant role in the determination of bank performance. In other words, the comparison of banks operating in different countries against a single reference may consider as inefficiency specific characteristics that a particular banking system is subject to and not whether its management of resources is effective. Dietsch and Lozano-Vivas (2000) show how different results may be if one controls for environmental conditions in the country where the bank is operating in relation to an uncontrolled specification. They state, therefore, that the estimation of one single frontier for heterogeneous banking markets may bias the results if one do not control for these factors.

Several papers have used the same method as Dietsch and Lozano-Vivas (2000) to deal with cross-country heterogeneity. Carvallo and Kasman (2005), for example, have also employed environmental variables in the estimation of cost inefficiency of Latin American countries. They include variables of three different types of classification: macroeconomic, banking system's and banking services accessability. Fang et al. (2011) utilize the GDP growth and inflation to differentiate banks from six transition South-Eastern European countries, while Lozano-Vivas and Pasiouras (2010) prefer to employ dummies that separate 87 countries in terms of their economic development. Common to all these papers is the recognition of the necessity in employing such variables in the frontier estimation.

Another solution for the problem of banking markets' heterogeneity is the

[1] See Berger and Humphrey (1997) for more information on bank efficiency estimation methods. This paper shows how even within a specific econometric model, one can vary the results due to the introduction or negligence of certain variables.

use of the fixed-effects estimator in the SFA. This method aims at removing any time invariant effect (observable or not) that may bias the inefficiency estimation. However, as Greene (2005) shows, the traditional fixed effects estimator does not separate efficiently the inefficient term and the individual heterogeneity. Therefore, he proposes a true fixed effects estimator, which is also expanded by Wang and Ho (2010). Both papers admit, however, that this method is prone to bias due to the incidental parameter problem as N tends to infinity.

Notwithstanding these couple of methods, we propose a new suitable model in order to explicitly model environmental factors in the estimation of technical efficiency. We consider that banks that are close geographically from one another are subject to similar constraints[2]. This approach is known as the geographically weighted stochastic frontier (GWSF), where we estimate frontiers for each bank in the sample (local frontiers). In each one we consider a different bank as a benchmark and a weight is given to other banks depending on the distance to this reference. This way we implicitly control for the geographic factors that influence efficiency of banks that are close to one another. An additional advantage of the GWSF is that we are able to employ it even within a country, in this case the US. Even though this geographic method has been applied in other papers, such as Samaha and Kamakura (2008) regarding the real state market, we are the first to employ it in panel data.

Despite the existence of some articles on this matter, the influence of geographic factors on banks' performance has not been properly recognized yet. In fact, this paper will show that there may be a significant bias in the efficiency scores if one does not take into account the geographical characteristics where each bank (or branch) operates. Some factors that influence bank efficiency may vary with the geographic location. Some are observable, such as: the size of the market, the different laws and regulations, the accessability of banking services by the population; others are unobservable. This method takes into account both types, since it estimates efficiency of a bank in comparison to its banks neighbors.

As a proof of the statement above, there is extensive evidence that US banks' performance is geographical dependent. Akhigbe and McNulty (2003) find that US commercial banks operating in metropolitan areas (MSA) have

[2]The literature has been interested in discussing whether recent technological developments, such as the more frequent use of internet and mobile banking, have reduced the importance of the physical location where bank operates. Even though internet plays an increasingly important role in reducing the costs of distance (Berger, 2003), Degryse and Ongena (2004) reaffirm the importance of the geographical distance in the lending relations.

different efficiency levels in relation to those in non-metropolitan areas for the years 1990-96. In fact, banks in MSA are less profit efficient than those in non-MSA. In addition, according to Tirtiroglu et al. (2011), bank productivity in the US appears to be geographically dependent amongst states, where performance in one state is positively correlated with the ones of its neighbors. Finally, Berger and DeYoung (2001) prove that return on assets varies considerably with the region. These facts are clear motivations for our exercise, where we apply this new method to US saving banks.

There is no denying that the study of US saving banks' efficiency fits our model, since these banks have a stronger regional focus of operation than regular commercial banks. In other words, they tend to lend more to institutions and enterprises that are close to where it is located[3]. US saving banks tend to compete with others that operate in the same geographic location, as well. It is, therefore, less likely that distant banks affect how a specific saving bank performs. Not only that, but these banks are also of uttermost importance due to the lending to small and medium firms. It is clearly the interest of bank's regulators to know exactly how these banks perform so as to choose the proper set of regulations for them.

In addition, the US banking system presents others interesting features on the geographic field. First, not only are these banks subject to a federal regulation, but they also have to respond to state laws, which may exert different influences in the banking operation. Second, as DeYoung et al. (2004) state[4], the removal of geographic restrictions that were in place since the McFadden Act of 1924 allowed banks to operate across state lines and to acquire banks anywhere in the country, converting some subsidiaries and removing branching restrictions. The Riegle-Neal Act of 1994, therefore, led

[3]Still regarding the topic of banking geography, there is a vast literature which highlight that the larger the distance between the bank and its clients, the less sustainable are its loans (see Moulton, 2010, for evidence in the mortgage market). In fact, Meyer and Yeager (2001) state that geographically concentrated banks are not vulnerable to downturns, even though their loan portfolio present a higher exposure to few sectors. The reason is that these banks develop personal relationships with clients that permit a more effective loan monitoring.

[4]DeYoung et al. (2004) considered these changes in economic conditions and explored whether effects in geography changed the bank headquarters locations, the bank branch office locations and the bank depositor locations. They found that (1) mergers and acquisitions have allowed banks to move headquarters from smaller to larger cities, (2) bank branches have moved farther away from headquarters and (3) spatial density of deposits in the 50 largest metropolitan areas has remained remarkably stable, since commercial banking industry became more spatially concentrated during the 1990s, as an evidence of gradual urbanization. The results suggest that spatial distribution of deposits remained similar across time. They also suggest that new technologies increase the ability of banks to manage credit relationships.

to a geographic expansion into new markets, where merger activities became more accepted by the banking industry. This merger process has increased and improved bank's ability to lend and to monitor these loans far away from headquarters. In fact, between 1980 and 1990, a period of consolidation and restructuring, banks were taken over by other depository institutions in order to raise efficiency.

We structure the remainder of the paper as follows. Section 2 presents our methodology, where we define the GWSF model and all the steps in order to estimate it. In Section 3, we show and summarize our the data sources. In addition, in Section 4, we present the empirical results, where we apply the GWSF to the case of US saving banks ad compare to it a fixed-effects specification. Finally, we make our concluding remarks in Section 5.

2 Methodology

In this study, we employ two different specifications of the stochastic frontier model (SF). One is the standard method that we estimate using fixed effects. In the other, we use geographically weighted estimation process (GWE), in which distinct coefficients are estimated for each bank. The GWE process, therefore, has the purpose of describing the effects of the regional environment over the functioning of US saving banks. Our purpose is that we may compare the results from these two specifications and finally reach relevant conclusions regarding the usefulness of the GWE in the efficiency estimation.

The basic SF model assumes that the production of a producer unit (company, government, machine, etc.) depends on the level of usage of required inputs, of a normal random shock (and other uncontrollable factors) that affects the productivity of the unit and of other component associated with the inefficiency of the unit, under managerial control. The latter component always takes positive values and therefore it must be associated to a strictly positive distribution. The degree of efficiency represents how close a bank is in relation to the stochastic frontier.

Thus, the SF model could be described, in its Cobb-Douglas version, as follows[5]:

$$AY_i = e^{\alpha_0} \left(\prod_{j=1}^{J} X_{ij}^{\alpha_j} \right) e^{\nu_i - \mu_i} \tag{1}$$

[5]For our purposes, the Cobb-Douglas version is more appropriate since it allows a more direct assessment to the elasticity of substitution of inputs and outputs as well as a clearer evaluation of the geographical intensity of these variables.

where AY_i is the output generated by producer i, X_{ij} is the utilization level of input j by producer i, the α's are coefficients to be estimated, ν_i is idiosyncratic term log-normally distributed with zero mean and standard deviation σ_ν, i.e.,$\nu_{it} \overset{iid}{\sim} N(0, \sigma_\nu^2)$ and μ_I is the inefficiency component with distribution log-normal truncated in one and standard deviation σ_μ, i.e. , i.e., $\mu_{it} \sim N^+(\overline{\mu}, \sigma_\mu^2)$.

Once this index has been achieved, we can use it in the production of several distinct outputs and thus, in the case of the Cobb-Douglas function:

$$AY_i = e^{\beta_0} \prod_{k=1}^{K} Y_{ik}^{\beta_k} \tag{2}$$

Where AY_i denotes the amount of output j produced by unit i. From these two equations one may extract the efficiency level of each unit:

$$\theta_i = e^{\mu_i} = e^{\alpha_0 - \beta_0} \prod_{j=1}^{J} X_{ij}^{\alpha_j} \prod_{k=1}^{K} Y_{ik}^{-\beta_k} \tag{3}$$

Here, θ_i is the efficiency level of unit i. As in the Cobb-Douglas case the homogeneity constraint over inputs holds if we normalize the variables in relation to one input. One can then write:

$$\frac{\theta_i}{X_{ij}} = e^{\gamma_0} X_{ij}^{-1} \prod_{j=1}^{J} X_{ij}^{\alpha_j} \prod_{k=1}^{K} Y_{ik}^{-\beta_k} = e^{\gamma_0} X_{ij}^{-1+\Sigma_{m=1}^{j}\alpha_m} \prod_{j=1}^{J} \left(\frac{X_{ij}}{X_{ik}}\right)^{\alpha_j} \prod_{k=1}^{K} Y_{ik}^{-\beta_k} \tag{4}$$

This, in its logarithmic version could be expressed as:

$$\log \theta_i - (\Sigma_m \alpha_m)\, x_{ij} = \gamma_0 + \sum_{j \neq J}^{J} \alpha_j x_{ij}^* - \sum_{k=1}^{K} \beta_k y_{ik} \tag{5}$$

Where, x and y are the neperian logarithms of X and Y, respectively, $x_j^* = \log \frac{X_j}{X_J}$ and $\gamma_0 = \alpha_0 + \beta_0$. This equation may be rearranged in order to isolate input x_i:

$$-x_{iJ} = \gamma_0 + \sum_{j \neq J}^{J} \left(\frac{\alpha_j}{\Sigma_m \alpha_m}\right) x_{ij}^* - \sum_{k=1}^{K} \left(\frac{\beta_k}{\Sigma_m \alpha_m}\right) y_{ik} - \log \theta_i \tag{6}$$

The next step is to interpret $-\log \theta_i$ as a residual, and thus one may utilize stochastic frontier techniques in order to estimate this input distance function. Thus:

$$-x_{iJ} = \gamma_0 + \sum_{j \neq J}^{J} \left(\frac{\alpha_j}{\Sigma_m \alpha_m} \right) x_{ij}^* - \sum_{k=1}^{K} \left(\frac{\beta_k}{\Sigma_m \alpha_m} \right) y_{ik} + \nu_i - \mu_i \tag{7}$$

with: $\log \theta_i = \mu_i - \nu_i$.

The parameters α are expected to be positive since that a reduction in the utilization of the reference input x_J will have to be compensated by an increase in the utilization of the other inputs. Conversely, parameters β are expected to be negative since that, a reduction in the utilization of the reference input will cause a decrease in the production everything else equal.

Taking this Cobb-Douglas function, the marginal rate of technical substitution between two inputs is given by:

$$MRTS_{i,j} = \frac{\alpha_j}{\alpha_i} \frac{X_i}{X_j} \tag{8}$$

It depends on the proportion in which the two inputs are being utilized. The ratio $\frac{\alpha_j}{\alpha_i}$ refers to the case in which both inputs are being utilized with same intensity. This ratio can be recovered from the estimated values in equation (5) by dividing these figures by each other in pairs.

In the GWE, we apply the maximum likelihood method sequentially to each unit, and each separate observation gains a weight according to the distance geographical relation to the reference unit. We assign these weights according to the following rule:

$$W_{ij} = \frac{e^{-\sqrt{\frac{d_{ij}}{\lambda}}}}{\sqrt{2\Pi\lambda}} \tag{9}$$

where W_{ij} is the weight of the j-unit in the estimation referenced over the i-unit, d_{it} is the great-circle distance in kilometers between the two units, λ is a dispersion parameter (bandwidth). If there are I units, we may then the normalize weights as follows:

$$\varpi_{ij} = \frac{IW_{ij}}{\sum_{k=1}^{J} W_{ik}} \tag{10}$$

In each estimation, the normalized weights are multiplied by their respective observations. As all units are used as reference by their turn, I estimations are performed and I sets of parameters are estimated, one for each unit.

The next step is to choose the appropriate λ. This parameter sets the weight distribution: the larger its magnitude, the greater the weight allocated to units farther away. The selection process is interactive and first we should establish a start value for it. In the algorithm created for this purpose the standard deviation of the distances between the units was used as a point of departure. It then proceeds to estimate the geographically weighted stochastic frontier (GWSF) and to collect the mean sum of squared residuals of the regressions obtained estimation process, which is the parameter to be minimized. The process is repeated with incremental variations in the bandwidth until the mean sum of squared residuals cease to decline, i.e. reaches it minimum.

When panel data is available, there are several different methods for estimating the inefficiency of the various producing units, see Kumbhakar and Lovell (2000) for a survey of these methods.

As the fixed-effects model allows for correlation between the regressors and the inefficiency component and between the latter and the idiosyncratic shock it seems to be the natural choice for us, since the input distance is formed by a composition of these two stochastic terms $-\ln(D_{Ii}) = \mu_i - \nu_i$. In panel data model, the equation to be estimated is:

$$-x_{iJ} = \gamma_0 + \sum_{j \neq J}^{J} \left(\frac{\alpha_j}{\Sigma_m \alpha_m} \right) x^*_{it,j} - \sum_{k=1}^{K} \left(\frac{\beta_k}{\Sigma_m \alpha_m} \right) y_{it,k} + \nu_i - \mu_i \tag{11}$$

where the subscript t refers to time. This equation may be modified to:

$$-x_{iJ} = \gamma_{i0} + \sum_{j \neq J}^{J} \left(\frac{\alpha_j}{\Sigma_m \alpha_m} \right) x^*_{it,j} - \sum_{k=1}^{K} \left(\frac{\beta_k}{\Sigma_m \alpha_m} \right) y_{it,k} + \nu_i \tag{12}$$

$$-x_{iKt} = \alpha_{0i} + \sum_{j \neq K}^{J} \alpha_j x^*_{ijt} + \sum_{j=1}^{J} \beta_j y_{ijt} + \nu_{it} \tag{13}$$

with $\gamma_{01} = \gamma_0 + \mu_i$, and then estimated by OLS. In order to find γ_0 and consequently μ_i, one could use the following estimators:

$$\hat{\gamma}_0 = \max_i(\hat{\gamma}_{i0}) \tag{14}$$

$$\hat{\mu}_i = \hat{\gamma}_0 - \hat{\gamma}_{0j} \tag{15}$$

And the technical efficiency of each unit can then be obtained with:

$$TE_i = e^{\hat{\mu}_i} \tag{16}$$

In the GWSF, we will have I sets of technical efficiencies, one for each weighted regression. In this case, the estimated inefficiency of one particular producing unit will be that obtained in the regression in which this specific unit was used as reference for the weights calculation.

3 Data Base and Variables

The estimations performed in this paper were based in an unbalanced panel, which contains registers of two hundred savings banks during nine years (2001-2009), totalizing 1260 observations. This database comprises two antagonic periods: one of financial stability until 2007 and another of financial turmoil after 2008. This fact allow us to draw interesting conclusions on how efficiency had evolved throughout these years. Savings banks of the sample are spread by 43 states and 172 towns. Map 1 shows how these banks are dispersed across US states.

[Map 1]

As the panel is unbalanced, most of the banks are not present in every year of the analysis due to the beginning/ending of operations or to missing data. Table 1 shows the quantity of banks present in the data set each year.

[Table 1]

The chosen variables for this exercise were the following: personnel expenses (input), interest expenses (input), other expenses (input, defined as total expenses minus personnel expenses minus interest expenses), bank's loans (output), liquid assets (output), total deposits (output), and non-interest income (output). Personnel expenses and interest expenses are usual measures of the banks' cost and, therefore, commonly employed as input variables. Other expenses, as said above, are the remaining expenses taken together. This split of the expenses is utilized because there may be an optimal combination of them, which enhance productivity. Most of the output variables are also quite traditional. Bank's loans, liquid assets and total deposits have been vastly utilized in the literature of banks' efficiency. The inclusion of non-interest income aims to capture the non-traditional bank activities, which are supposed to be quite distinct in geographical terms[6].Even recognizing that other variables may be geographically dependent, we believe

[6]Even though there is not a consensus on the matter yet, the literature has given an increasing importance in incorporating variables of bank non-traditional activities (such as off-balance sheet and non-interest income) in the analysis of bank efficiency (Lozano-Vivas and Pasiouras, 2010). Ignoring these measures can be misleading, since it does not take into account the bank's balance sheet as a whole

that the introduction of the non-interest income might be more appropriate to capture this aspect.

[Table 2]

Table 2 presents some descriptive statistics of the utilized variables. Loans and deposits are the main outputs of US saving banks as this table shows. As expected, the most important input is interest expenses, since banks main activity is to intermediate interest-bearing funds. On the other hand, personnel expenses account for approximately 18% of the expenses on average. This proportion can be considered high and it evidentiates that saving banks are more labor intensive than regular commercial banks. Finally, one can observe that all variables have a high standard deviation, which shows that we consider very heterogeneous banks in our specifications.

4 Empirical Results

This section presents our model's empirical results. First, as we have already stated in section 2, we estimate an input-distance function using other bank' expenses as reference (endogenous) variable. We perform this estimation in a non-spatial context by fixed-effect OLS. Our choice for reference input was based in the fact that, other bank' expenses is a variable which may include several distinct components and thus, the analysis of its coefficient would not be very informative.

We also include a time trend and its square in the estimated equation in order to capture the temporal tendency of the banks' efficiency. Thus, the estimated equation can be written as:

$$-x_{it,1} = \alpha_1 + \alpha_2 x*_{it,2} + \alpha_3 x*_{it,3} + \beta_1 y_{it,1} + \beta_2 y_{it,2} + \beta_3 y_{it,3} + \beta_4 y_{it,4} + \beta_t t + \beta_T t^2 + \nu_{it} + \mu_{it} \quad (17)$$

where $x_{it,1}$ is the logarithm of the other bank's expenses , $x*_{it,2}$ is the logarithm of the ratio between personnel expenses and other bank's expenses, $x*_{it,3}$ is the logarithm of the ratio between interest expenses and other bank's expenses, $y_{it,1}$ is the logarithm of the bank's loans, $y_{it,2}$ is the logarithm of the total liquid assets, $y_{it,3}$ is the logarithm of the total deposits and $y_{it,4}$ is the logarithm of the non-interest income. The use of a non-traditional activity output is necessary to avoid bias in the results. According to Lozano-Vivas and Pasiouras (2010), these activities have been increasingly important in the bank's balance sheet. t and t^2 are, respectively, the year of the observation and its square. Table 3 shows the results of the estimated equation.

[Table 3]

The values of the R^2 and of the F-statistic are satisfactory, and indicate a good fit of the model. The t-statistics show significance of the parameters for all variables, except total liquid assets and the time trend in level. We will provide a further analysis on the interpretation of this model's coefficients next. The purpose is to determine how each variable affects the employment of the reference input. Then, we perform a comparison of the fixed effects estimator against the geographically weighted model.

It is clear that all input and output coefficients possess the previously expected signs. The input variables' positive values mean that, the greater utilization of any input, the smaller the necessity of the utilization of the reference input[7]. In other words, given a determined amount of bank input, the use of one additional unit of input X_1 means the lower employment of inputs X_i, $\forall\ i \neq 1$ in one unit. Reciprocally, the greater the production of any output, the greater the utilization of the reference input (everything else constant). It is reasonable to suppose that the production of one additional unit of output requires more inputs in general.

Another inference from the values of the coefficients is that US saving banks seem to have increasing returns of scale. The scale elasticity is equal to the ratio between the sum of the coefficients of the input variables, plus one, and the sum of the absolute values of the coefficients of the output variables. In this case, the estimated scale elasticity for the non-spatial model was 2.18. This result implies that US saving banks have not yet achieved their optimal size in terms of technical efficiency.

The coefficient associated to the time trend is positive, but not significative, while that linked to its square is negative and significant. This can mean either that there is a negative evolution of the banks' efficiency, or that the coefficient of the time trend is truly positive, but the estimation fails in recognizing this fact. Fortunately, the results of the GWE suggest that the latter possibility is probably true.

Additionally, a traditional Translog transformation function was also estimated in order to test for model robustness. The correlation between Cobb-Douglas and Translog results was 0.79, which indicates a good adherence between the two models results. The Cobb-Douglas estimation pointed for an average efficiency of 0.462, while the Translog model produced an efficiency mean of 0.350.

The second estimation that we perform is the GWE. This process involves 150 sub-estimations, each of them with a distinct set of weights. One can

[7]Remembering: the reference variable is taken in negative values.

obtain this change of the weights by varying the standard deviation of the normal generating function. We began with a standard deviation of 100 km and went up to 15,000 km. The better parameter, measured by the minimum average of the square residuals sum, is equal to 2,400km. In graph 1, we show the relation between the standard deviation of the weight-generating function (lamb) and the exponential of the average sum of the square residuals.

Once more a translog model was also tested, this time with geographical weights. The comparison with results of the Cobb-Douglas estimation, also point to the model robustness. The correlation between banks technical efficiency obtained for the two methods was 0.85, and the correlation between efficiency ranks was 0.9.

[Graph 1]

Table 4 pictures the results of the GW estimation. This table presents some descriptive statistics of the coefficients that we have estimated in all the local sub-estimations. In order to clarify any possible confusion, it is worthwhile to remark that the standard deviations in the third row are not the parameter associated with the estimator, but the deviations of the estimated values. The last row once more shows the results of the non-spatial estimation in order to facilitate the comparison. One remarkable fact is that in the GW estimation all coefficients turned up significant.

[Table 4]

All estimations also provide evidence of increasing returns to scale, although somewhat smaller that the non-spatial estimation obtains (1.98 against 2.18). This means that the fixed effects model has super-estimated the aggregate effects of inputs on the reference and/or sub-estimated the impact of output on the bank's other expenses. The optimal size of US saving banks is somewhat lower when we control for geographic factors.

GWE also clarifies the time trend of the efficiency, which is increasing but with diminishing returns. That is to say that in the last years of the sample, the trend's inclination has decreased. This is consistent with the occurrence of the financial crisis, that has directly affected the performance of US banks.

Comparing the estimated coefficient of both spatial and non-spatial frontiers, some interesting insights may be drawn. First, the coefficient associated to personnel expenses (α_2) is always greater in the non-spatial estimation in relation to the GWE. Map 2 shows that α_2 varies from 0.453 to 0.461, where in the non-spatial case it equals 0.545. This implies a greater marginal rate of technical substitution between personnel expenses and other expenses in the former than in the latter.

[Map 2]

The inverse seems to occur with the interest expenses. As one can observe in Map 3, their GW coefficients have always greater values than that of the non-spatial case.

[Map 3]

The significance of these results is that, once you control for geographic factors, it is easier to substitute manpower by other kind of inputs, but it is harder to do the same about loaned funds.

In terms of outputs, the major distinction that emerged was about the total deposits, this item seems to be much more relevant in the spatial case. The larger coefficient associated to this variable implies that a larger foregone of deposits is necessary in order to generate a given amount of other outputs at the regional level. This effect points out to the importance of local branches in the capture of deposits.

[Map 4]

Non-interest income, contrary to expected, did not show great distinction between spatial and non-spatial models. Indeed, the non-spatial case showed a slightly greater coefficient, although the difference between them is not significant.

[Map 5]

When technical efficiencies from the two models are contrasted, one can observe that 3/4 (126) of the banks had their efficiency improved in the spatial model. This phenomena has occurred because that model compares banks compares to others near by, which probably are subject to similar opportunities and constraints. One direct implication is that comparisons are more flexible in this model. That is, a bank's efficiency is not measured against a standard established by faraway institutions of very distinct nature and with a different environment. Therefore, a bank that is not considered efficient in comparison to the best practice bank in the sample may be very efficient in relation to those in its vicinity.

Spite of this large proportion of banks that had their efficiency enhanced in the spatial model, there were no substantial changes in efficiency average or standard deviation from one model to another. This means that, once we define a spatial model, efficiency was, to a large extent, redistributed from those who presented a higher performance in the FE model to those with lower scores.

[Map 6]

[Map 7]

A comparison between maps 6 and 7 permits one to analyze the differences in efficiency in the spacial and non-spatial models. It is clear that banks from the Northeast of the US appear to be less efficient in the GWSF model. To the contrary, the West and South of the country have banks with a higher efficiency in the spatial case.

5 Conclusions

This paper has applied a geographically weighted stochastic frontier model to a panel data in order to determine the efficiency levels of US saving banks between 2001 and 2009. Since there is a large evidence that environment matters when estimating efficiency frontiers, we propose to control it through the GWSF method. Besides, taking into account the regulations and economic environment of different locations in the estimation of the frontier, it has never been employed in panel data until now, specially in the banking literature. In order to show the advantages and viability of this model, we compare these results with a fixed-effects SF model.

Many studies have applied the regular SF model and controlled for unobservable factors using macroeconomic variables or even a fixed effects model. The problem of the former is related to the choice of variables to employ in the specification. In other words, one may not be certain about which macroeconomic factors have a significant effect on bank efficiency. In addition, there may be several unobservable variables that do not have a suitable proxy. We have shown that the latter, on the other hand, by eliminating from the specification time-invariant factors, may also sub-estimate efficiency by itself. Also, environmental characteristics that have changed in a particular period of time would not be captured by the fixed effect approach.

The case of US saving banks, the subject of our analysis, is evidence to support GWSF approach. These banks operate in a regional level by lending to small and medium enterprises and, therefore, local characteristics may have a higher influence in their behavior than a bank that operates in the whole country. When we control for geographic factors, the technical efficiency appear to be larger for most of the banks, since the ones that operate close to a specific bank are given a higher weight in the estimation. In other words, the bank performance is now compared to those that are subject to the same constraints and not to those that have completely different conditions. Our overall conclusion is that geography matters and it plays an essential role in correctly estimating efficiency. This result has important implications for policy makers, since one policy does not necessarily fits all.

Amongst the secondary findings, we highlight some of them next: (i) we find in the GWSF model an positive time trend of efficiency for the period. This positive trend, however, decreases as the period we consider is near to the end. This is an evidence of the financial turmoil effects on US bank efficiency that the FE model has failed to capture; (ii) there has been differences between both models in terms of both inputs and outputs' estimated coefficients, as well, with the exception of non-interest income. The substitution between this last one and the reference input appears to be insensitive to local conditions and constraints.

Further analysis could employ the geographic weighted method to the estimation of other important parameters to the banking literature. One in particular may be competition, since banks may compete more with banks that are close geographically in relation to those that are more distant.

References

Aigbe Akhigbe and James E. McNulty. The profit efficiency of small US commercial banks. *Journal of Banking and Finance*, 27:307325, 2003.

A. N. Berger. The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, Credit and Banking*, 35:141–176, 2003.

A. N. Berger and R. DeYoung. The effects of geographic expansion on bank efficiency. *Journal of Financial Services Research*, 19:163–184, 2001.

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

A. N. Berger, I. Hasan, and M. Zhou. Bank ownership and efficiency in China: What will happen in the world's largest nation? *Journal of Banking & Finance*, 33:113–130, 2009.

O. Carvallo and A. Kasman. Cost efficiency in the Latin American and Caribbean banking systems. *Journal of International Financial Markets, Institutions and Money*, 15:55–72, 2005.

Hans Degryse and S. Ongena. The impact of technology and regulation on the geographical scope of banking. *Oxford Review of Economic Policy*, 20 (4):571–590, 2004.

Robert DeYoung, Thomas Klier, and Daniel P. McMillen. The changing geography of the U.S. banking industry. *The Industrial Geographer*, 2: 29–48, 2004.

Michel Dietsch and Ana Lozano-Vivas. How the environment determines banking effciency: A comparison between French and Spanish industries. *Journal of Banking and Finance*, 24:985–1004, 2000.

Y. Fang, I. Hasan, and K. Marton. Bank efficiency in transition economies: Recent evidence from South-Eastern Europe. Research Discussion Papers 5/2011, Bank of Finland, March 2011.

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

S. C. Kumbhakar and C.A.K. Lovell. Stochastic frontier analysis. *European Journal of Operational Research*, 2000.

Ana Lozano-Vivas and Fotios Pasiouras. The impact of non-traditional activities on the estimation of bank efficiency: International evidence. *Journal of Banking and Finance*, 34:1436–1449, 2010.

A. Meyer and T. Yeager. Are small rural banks vulnerable to local economic downturns? Review, Federal Reserve Bank of St. Louis, Mar 2001.

S. Moulton. Originating lender localness and mortgage sustainability: An evaluation of deliquency and foreclosure in Indiana's mortgage revenue bond program. *House Policy Debate*, 20(4):581–617, 2010.

Stephen A. Samaha and Wagner A. Kamakura. Assessing the market value of real state property with a geographically weighted stochastic frontier model. *Real Estate Economics*, 36(4):717–751, 2008.

P. L. Tecles and B. M. Tabak. Determinants of bank efficiency: The case of Brazil. *European Journal of Operational Research*, 207:1587–1598, 2010.

Dogan Tirtiroglu, A. Basak Tanyeri, Ercan Tirtiroglu, and Mehmet Kenneth N. Daniels. Banking geography and cross-fertilization in the productivity growth of us commercial banks. Working Papers 1108, TUSIAD-Koc University Economic Research Forum, April 2011.

H. Wang and C. Ho. Estimating fixed-effect panel stochastic frontier models by model transformation. *Journal of Econometrics*, 157:286–296, 2010.

Table 1: Number of Banks in the Data set by Year

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
Banks	81	198	198	170	162	152	135	122	109

Table 2: Descriptive Statistics of the Utilized Variables (in US$ millions)

Variables	Status	Mean	Median	St. Dev.	Max.	Min.
Personnel Expenses	Input	45.142	16.388	88.932	810.737	809
Interest Expenses	Input	120.028	31.050	289.004	4.710.007	342
Other Expenses	Input	90.183	15.594	356.884	6.087.496	555
Bank's Loans	Output	3.607.313	865.508	9.487.042	125.167.453	23.428
Liquid Assets	Output	158.349	45.010	455.762	7.068.700	1.164
Total Deposits	Output	2.895.914	920.398	6.117.170	69.603.422	41.658
Non-interest Income	Output	72.070	9.616	270.300	4.098.312	63

Table 3: Fixed-Effect, Non-spatial, OLS Estimation

Variable	$\mathbf{x}_2^*$	$\mathbf{x}_3^*$	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{y}_4$	$\mathbf{t}$	$\mathbf{t}^2$
Coefficient	0.545	0.237	-0.351	-0.011	-0.296	-0.159	0.008	-0.005
St Dev	0.021	0.017	0.033	0.009	0.035	0.011	0.01	0.001
t-stat	25.902	14.032	-10.793	-1.196	-8.331	-14.479	0.778	-4.579
p-value	0	0	0	0.232	0	0	0.437	0
$R^2 = 0.872$			F -stat = 899.762				F p-value = 0.000	

Table 4: Descriptive Statistics of the Estimated GWE Coefficients

Coefficient	α_2	α_3	β_1	β_2	β_3	β_4	β_t	β_T
Mean	0.459	0.282	-0.345	-0.027	-0.359	-0.144	0.02	-0.006
Median	0.46	0.281	-0.345	-0.027	-0.359	-0.145	0.02	-0.006
Stand. Dev	0.002	0.001	0.001	0.001	0.001	0.001	0	0
Max	0.461	0.285	-0.344	-0.027	-0.358	-0.142	0.021	-0.006
Min	0.453	0.281	-0.348	-0.029	-0.36	-0.145	0.02	-0.006
Non-spatial	0.545	0.237	-0.351	-0.011	-0.296	-0.159	0.008	-0.005

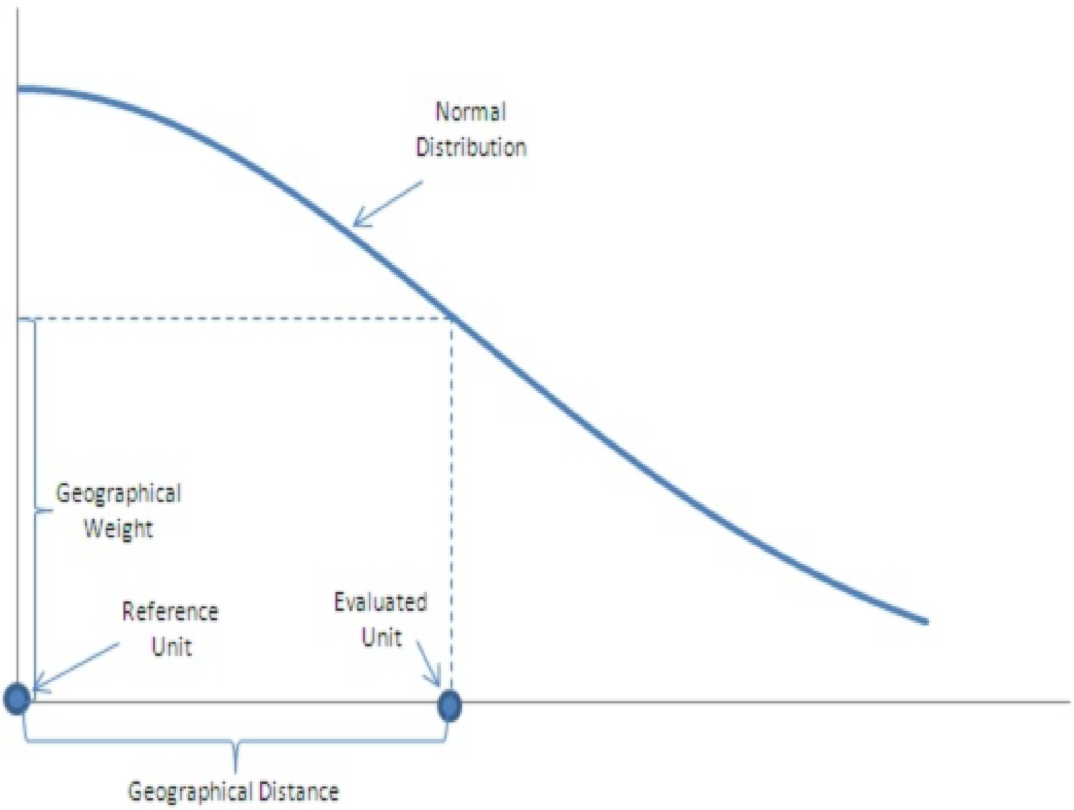


Figure 1: *Graph1*

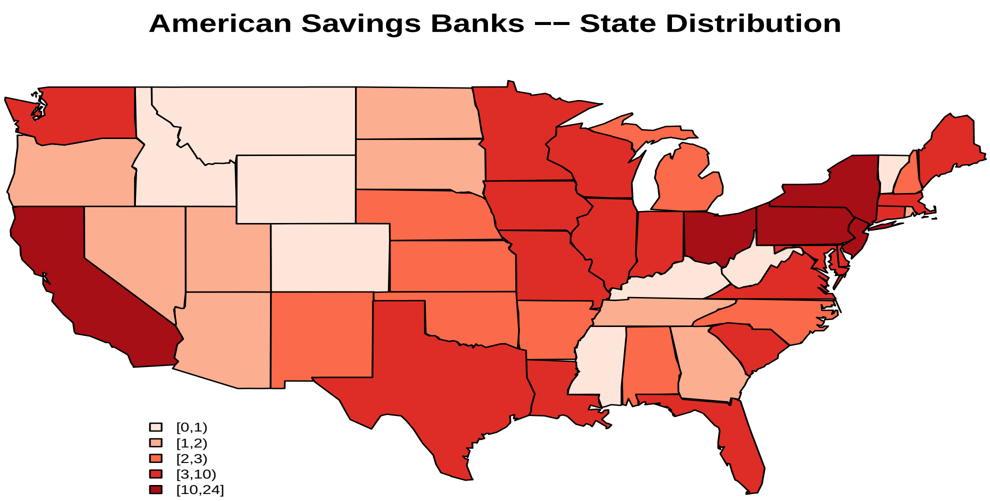


Figure 2: *Map1*

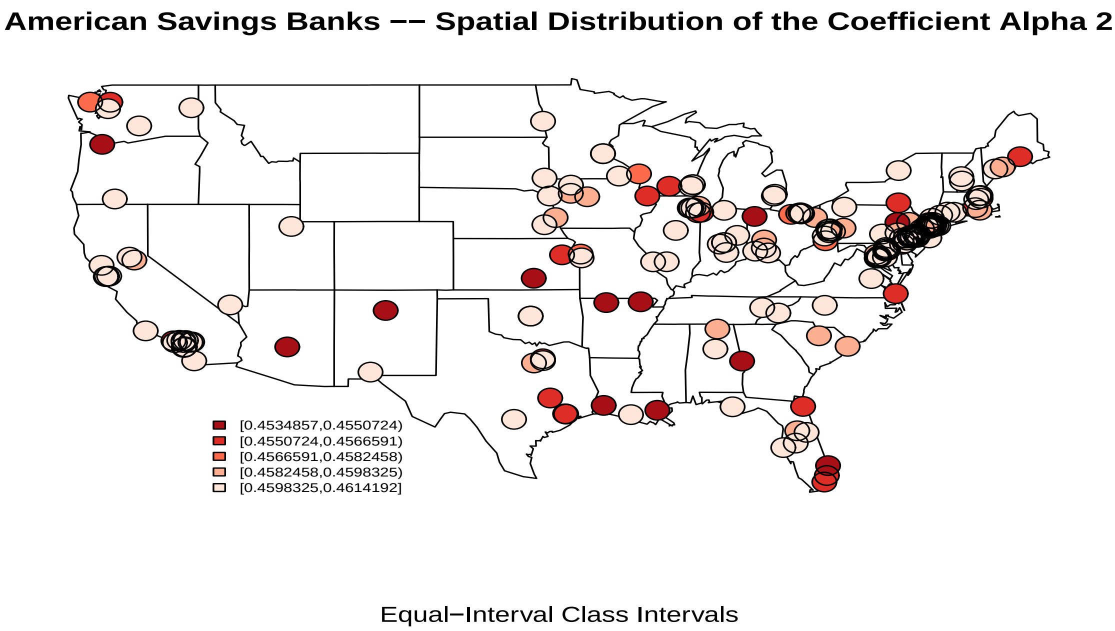
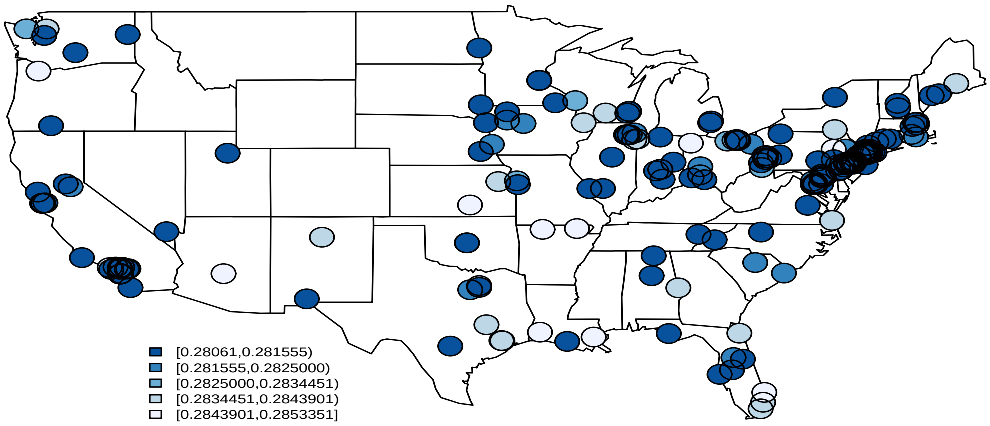


Figure 3: *Map2*

American Savings Banks −− Spatial Distribution of the Coefficient Alpha 3



Equal−Interval Class Intervals

Figure 4: *Map3*

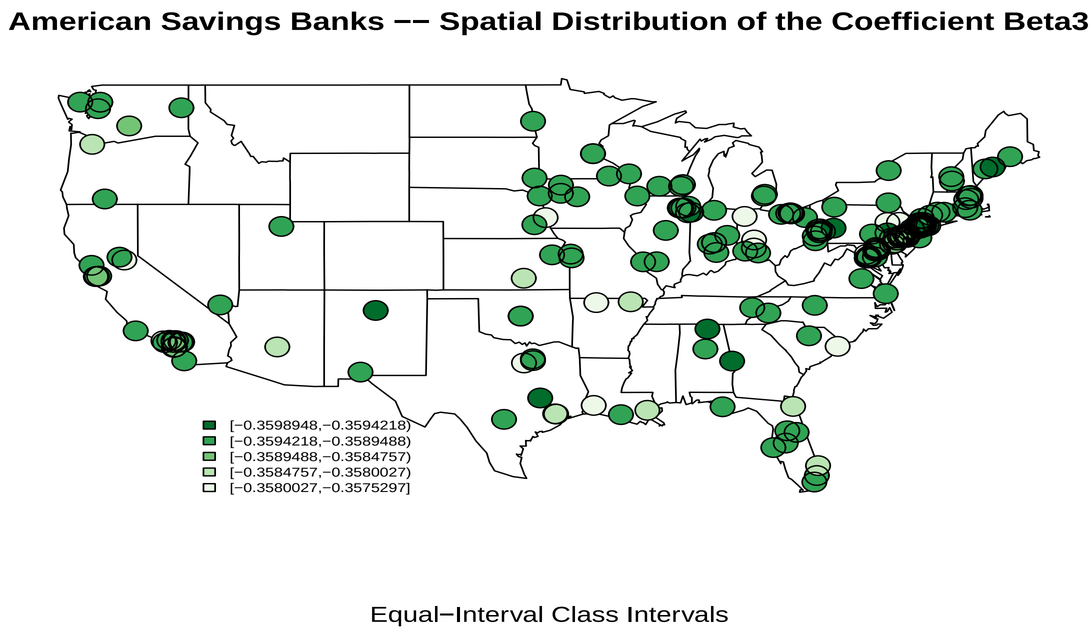


Figure 5: *Map4*

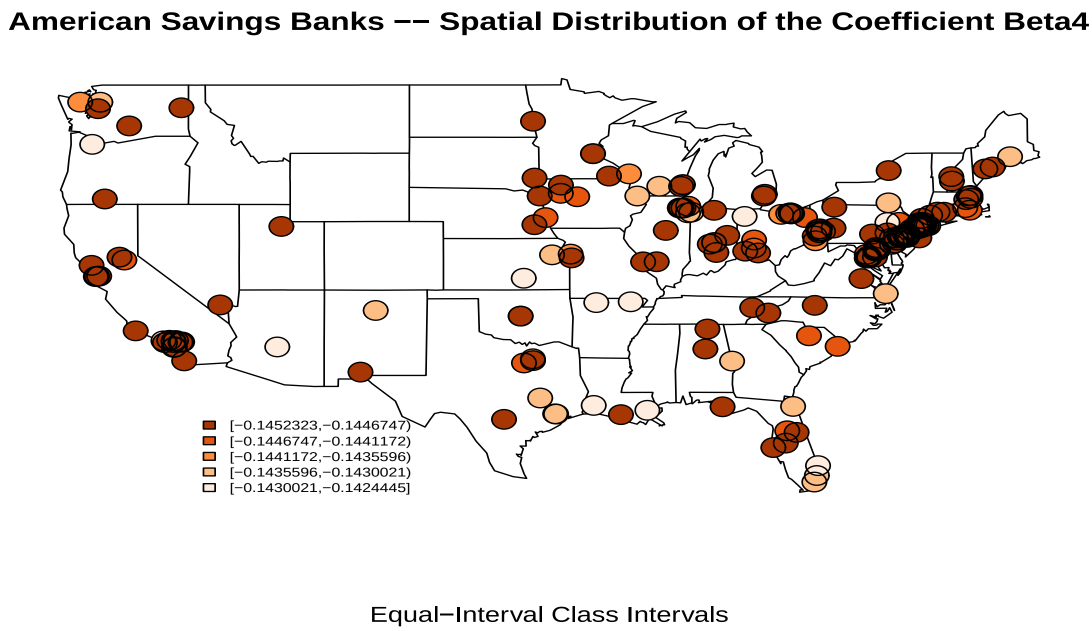


Figure 6: *Map5*

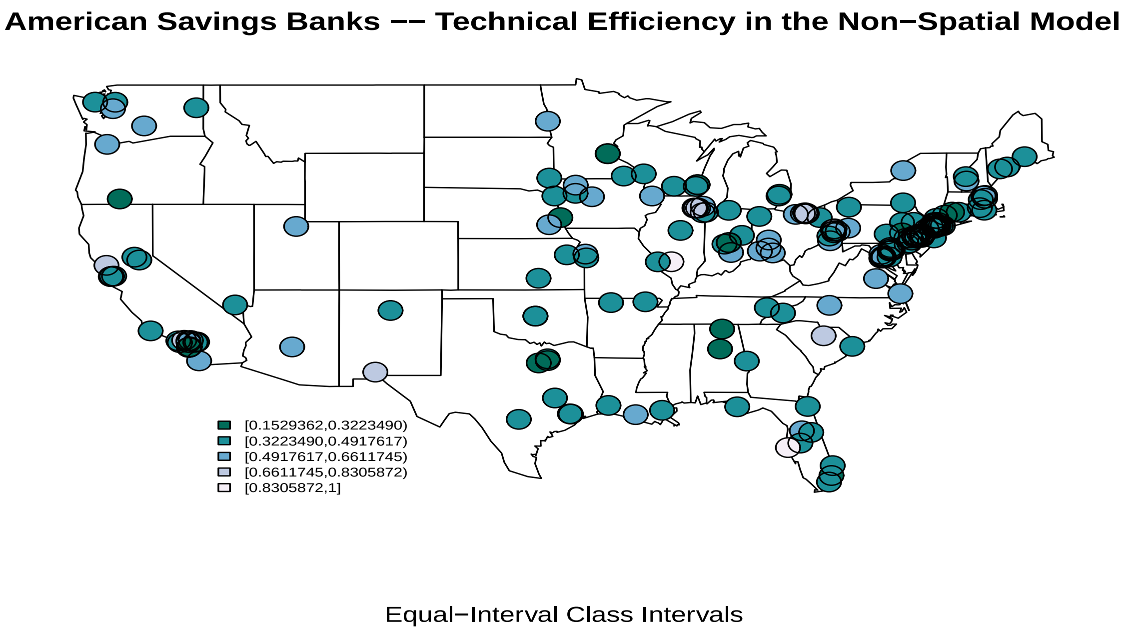


Figure 7: *Map6*

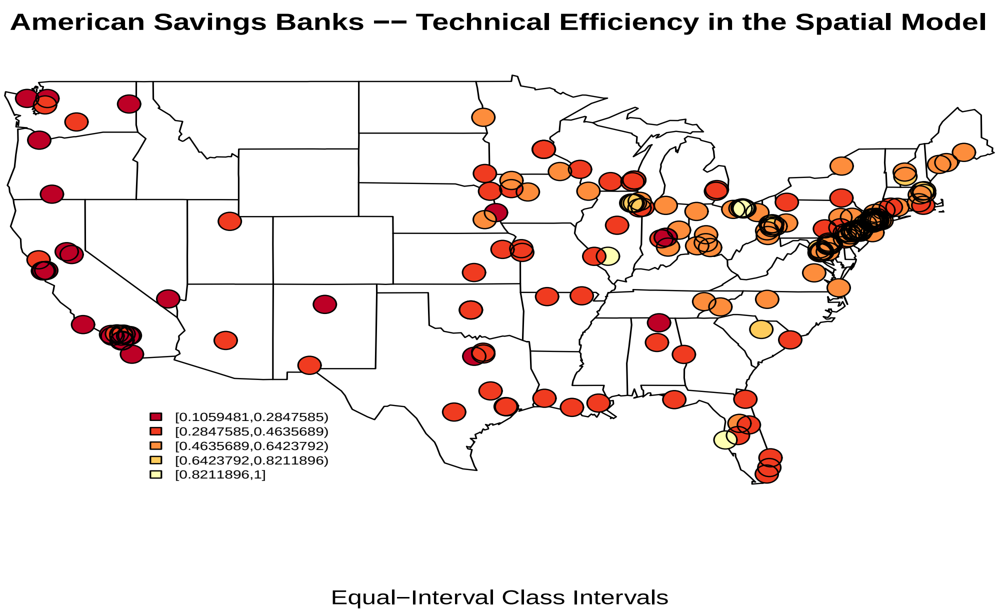


Figure 8: *Map7*

Banco Central do Brasil

Trabalhos para Discussão

Os Trabalhos para Discussão do Banco Central do Brasil estão disponíveis para download no website
http://www.bcb.gov.br/?TRABDISCLISTA

Working Paper Series

The Working Paper Series of the Central Bank of Brazil are available for download at
http://www.bcb.gov.br/?WORKINGPAPERS

236	**Optimal costs of sovereign default** *Leonardo Pio Perez*	Apr/2011
237	**Capital Regulation, Monetary Policy and Financial Stability** *P.R. Agénor, K. Alper, and L. Pereira da Silva*	Apr/2011
238	**Choques não Antecipados de Política Monetária e a Estrutura a Termo das Taxas de Juros no Brasil** *Fernando N. de Oliveira e Leonardo Ramos*	Abr/2011
239	**SAMBA: Stochastic Analytical Model with a Bayesian Approach** *Marcos R. de Castro, Solange N. Gouvea, André Minella, Rafael C. Santos and Nelson F. Souza-Sobrinho*	Apr/2011
240	**Fiscal Policy in Brazil through the Lens of an Estimated DSGE Model** *Fabia A. de Carvalho and Marcos Valli*	Apr/2011
241	**Macro Stress Testing of Credit Risk Focused on the Tails** *Ricardo Schechtman and Wagner Piazza Gaglianone*	May/2011
242	**Determinantes do *Spread* Bancário *Ex-Post* no Mercado Brasileiro** *José Alves Dantas, Otávio Ribeiro de Medeiros e Lúcio Rodrigues Capelletto*	Maio/2011
243	**Economic Activity and Financial Institutional Risk: an empirical analysis for the Brazilian banking industry** *Helder Ferreira de Mendonça, Délio José Cordeiro Galvão and Renato Falci Villela Loures*	May/2011
244	**Profit, Cost and Scale Eficiency for Latin American Banks: concentration-performance relationship** *Benjamin M. Tabak, Dimas M. Fazio and Daniel O. Cajueiro*	May/2011
245	**Pesquisa Trimestral de Condições de Crédito no Brasil** *Clodoaldo Aparecido Annibal e Sérgio Mikio Koyama*	Jun/2011
246	**Impacto do Sistema Cooperativo de Crédito na Eficiência do Sistema Financeiro Nacional** *Michel Alexandre da Silva*	Ago/2011
247	**Forecasting the Yield Curve for the Euro Region** *Benjamim M. Tabak, Daniel O. Cajueiro and Alexandre B. Sollaci*	Aug/2011

248	**Financial regulation and transparency of information: first steps on new land** *Helder Ferreira de Mendonça, Délio José Cordeiro Galvão and Renato Falci Villela Loures*	Aug/2011
249	**Directed clustering coefficient as a measure of systemic risk in complex banking networks** *B. M. Tabak, M. Takami, J. M. C. Rocha and D. O. Cajueiro*	Aug/2011
250	**Recolhimentos Compulsórios e o Crédito Bancário Brasileiro** *Paulo Evandro Dawid e Tony Takeda*	Ago/2011
251	**Um Exame sobre como os Bancos Ajustam seu Índice de Basileia no Brasil** *Leonardo S. Alencar*	Ago/2011
252	**Comparação da Eficiência de Custo para BRICs e América Latina** *Lycia M. G. Araujo, Guilherme M. R. Gomes, Solange M. Guerra e Benjamin M. Tabak*	Set/2011
253	**Bank Efficiency and Default in Brazil: causality tests** *Benjamin M. Tabak, Giovana L. Craveiro and Daniel O. Cajueiro*	Oct/2011
254	**Macroprudential Regulation and the Monetary Transmission Mechanism** *Pierre-Richard Agénor and Luiz A. Pereira da Silva*	Nov/2011
255	**An Empirical Analysis of the External Finance Premium of Public Non-Financial Corporations in Brazil** *Fernando N. de Oliveira and Alberto Ronchi Neto*	Nov/2011
256	**The Self-insurance Role of International Reserves and the 2008-2010 Crisis** *Antonio Francisco A. Silva Jr*	Nov/2011
257	**Cooperativas de Crédito: taxas de juros praticadas e fatores de viabilidade** *Clodoaldo Aparecido Annibal e Sérgio Mikio Koyama*	Nov/2011
258	**Bancos Oficiais e Crédito Direcionado – O que diferencia o mercado de crédito brasileiro?** *Eduardo Luis Lundberg*	Nov/2011
259	**The impact of monetary policy on the exchange rate: puzzling evidence from three emerging economies** *Emanuel Kohlscheen*	Nov/2011
260	**Credit Default and Business Cycles: an empirical investigation of Brazilian retail loans** *Arnildo da Silva Correa, Jaqueline Terra Moura Marins, Myrian Beatriz Eiras das Neves and Antonio Carlos Magalhães da Silva*	Nov/2011
261	**The relationship between banking market competition and risk-taking: do size and capitalization matter?** *Benjamin M. Tabak, Dimas M. Fazio and Daniel O. Cajueiro*	Nov/2011
262	**The Accuracy of Perturbation Methods to Solve Small Open Economy Models** *Angelo M. Fasolo*	Nov/2011

263 **The Adverse Selection Cost Component of the Spread of Brazilian Stocks** Nov/2011
Gustavo Silva Araújo, Claudio Henrique da Silveira Barbedo and José Valentim Machado Vicente

264 **Uma Breve Análise de Medidas Alternativas à Mediana na Pesquisa de Expectativas de Inflação do Banco Central do Brasil** Jan/2012
Fabia A. de Carvalho

265 **O Impacto da Comunicação do Banco Central do Brasil sobre o Mercado Financeiro** Jan/2012
Marcio Janot e Daniel El-Jaick de Souza Mota

266 **Are Core Inflation Directional Forecasts Informative?** Jan/2012
Tito Nícias Teixeira da Silva Filho

267 **Sudden Floods, Macroprudention Regulation and Stability in an Open Economy** Feb/2012
P.-R. Agénor, K. Alper and L. Pereira da Silva

268 **Optimal Capital Flow Taxes in Latin America** Mar/2012
João Barata Ribeiro Blanco Barroso

269 **Estimating Relative Risk Aversion, Risk-Neutral and Real-World Densities using Brazilian Real Currency Options** Mar/2012
José Renato Haas Ornelas, José Santiago Fajardo Barbachan and Aquiles Rocha de Farias

270 **Pricing-to-market by Brazilian Exporters: a panel cointegration approach** Mar/2012
João Barata Ribeiro Blanco Barroso

271 **Optimal Policy When the Inflation Target is not Optimal** Mar/2012
Sergio A. Lago Alves

272 **Determinantes da Estrutura de Capital das Empresas Brasileiras: uma abordagem em regressão quantílica** Mar/2012
Guilherme Resende Oliveira, Benjamin Miranda Tabak, José Guilherme de Lara Resende e Daniel Oliveira Cajueiro

273 **Order Flow and the Real: Indirect Evidence of the Effectiveness of Sterilized Interventions** Apr/2012
Emanuel Kohlscheen

274 **Monetary Policy, Asset Prices and Adaptive Learning** Apr/2012
Vicente da Gama Machado