Bayesian Inference on Technology and Cost Efficiency of Bank Branches*

Wnioskowanie bayesowskie o technologii i efektywności kosztowej oddziałów banku

Jerzy Marzec**, Jacek Osiewalski***

received: 29 April 2008, final version received: 12 September 2008, accepted: 8 October 2008

Abstract

We present Bayesian statistics and Gibbs sampling, an MCMC simulation technique, as tools for making inferences in stochastic frontier models for panel data from the banking sector. In our empirical example, the Bayesian approach is applied to estimate a short-run frontier cost function for $N = 58$ branches of a Polish commercial bank, observed over $T = 4$ quarters of one year. We use a translog cost function (with regularity conditions imposed for an 'average' branch) and treat inefficiency as a random individual effect, assuming a varying efficiency distribution (VED) specification proposed by Koop, Osiewalski and Steel (1997).

Keywords: Bayesian econometrics, panel data, cost models, microeconomics of bank.

JEL: C11, C23, D24, G21

Streszczenie

W artykule prezentujemy statystykę bayesowską i próbkowanie Gibbsa (technikę symulacji typu MCMC) jako narzędzia wnioskowania w stochastycznych modelach granicznych dla danych panelowych z sektora bankowego. W naszym przykładzie empirycznym podejście bayesowskie służy do estymacji krótkookresowej granicznej funkcji kosztu dla 58 oddziałów polskiego banku komercyjnego, na podstawie danych z 4 kwartałów jednego roku. Przyjmujemy funkcję kosztu typu translog (z warunkami regularności dla przeciętnego oddziału), a nieefektywność traktujemy jak losowy efekt indywidualny, wykorzystując specyfikację o zmiennym rozkładzie efektywności (VED), którą zaproponowali Koop, Osiewalski i Steel (1997).

Słowa kluczowe: ekonometria bayesowska, dane panelowe, modele kosztu, mikroekonomia banku.

* A much earlier version of this work was presented at the Wolpertinger Conference of the European Association of University Teachers of Banking and Finance (Cracow University of Economics, Kraków, 2-4 September, 2004).

** Cracow University of Economics, Department of Econometrics; e-mail: eemarzec@cyfronet.krakow.pl

*** Cracow University of Economics, Department of Econometrics; Andrzej Frycz Modrzewski Cracow University College; e-mail: eeosiewa@cyf-kr.edu.pl

1. Introduction

The stochastic frontier or composed error framework was first introduced in Meeusen and van den Broeck (1977) and Aigner et al. (1977) and has been used in many empirical applications. In particular, stochastic frontier models have been applied in studies of production and cost efficiency in the banking sector; see Ferrier, Lovell (1990), Cebenoyan et al. (1993), Bauer, Hancock (1993), Mester (1993; 1997), Berger, Mester (1997), Berger, De Young (1997), Kraft, Tirtiroglu (1997), Altunbas et al. (2000). All these empirical studies used the sampling theory (classical) methods of inference.[1]

Van den Broeck, Koop, Osiewalski and Steel (1994), hereafter BKOS, Koop, Steel and Osiewalski (1995), and Koop, Osiewalski and Steel, hereafter KOS (1994a; 1994b; 1997; 1999; 2000a; 2000b) used Bayesian methods to analyze stochastic frontier models and argued that such methods had several advantages over their classical counterparts in the treatment of these models. Most importantly, the Bayesian approach enables to provide exact finite sample results for any feature of interest and to take fully into account parameter uncertainty. The Bayesian statistical methodology has been successfully applied in various empirical issues, ranging from hospital efficiencies in KOS (1994b; 1997) to analyses of the growth of countries in KOS (1999; 2000a; 2000b). In this paper we apply the Bayesian approach to model the short-run cost frontier and to measure cost efficiency of bank branches.

There are different reasons for focusing on branches of one of Polish commercial banks; some reasons are practical and rather specific to the situation of the banking sector in Poland, other are more general and of methodological nature. First of all, it was much easier to collect (or, in fact, to construct – see Marzec 2000) reliable and fully comparable data representing activities of all branches of a big Polish bank than to find a data set of similar quality that would represent a relatively homogenous (and not too small) group of Polish banks. Thus, focusing on branches of one bank helped us to avoid problems with heterogeneity, discussed by Mester (1997). Second, branches (as opposed to specialized departments or units) are not involved in financial services that would be called "nontraditional activities". As regards the case we report in the empirical example, the branches under study represented traditional banking technology which can be modelled within the framework of Sealey and Lindley (1977). This means we could focus on the presentation of the Bayesian statistical methodology at work and not on addressing new questions related to the economics of a bank.[2] Third, modelling the technology used by branches and making efficiency comparisons among them constitute a very useful tool for the management of the bank. The analysis of activities of bank branches was presented by Zardokoohi and Kolari (1994), Berger et al. (1997) and others, using mainly mathematical programming techniques.

In our first preliminary study we used only cross-sectional data and a very simplified cost frontier; see Osiewalski, Marzec (1998). This work is based on a more mature approach, already adopted in our papers published only in Polish; see Marzec, Osiewalski (2001; 2003). Here we summarise and extend our previous research. Thus, we use panel data and a translog cost function. We show how inferences on technology and individual cost efficiencies of bank branches can be made using Bayesian random effects models proposed in KOS (1997) and a variant of the Gibbs sampler developed therein. We adopt the general Varying Efficiency Distribution (VED) model specification and apply a Highest Posterior Density (HPD) test to examine statistical validity of the simpler, nested CED (Common Efficiency Distribution) model. Our approach enables to impose (locally) all economic regularity conditions on the short-run translog cost model.

2. The Bayesian Stochastic Frontier Model

The basic sampling model considered here can be written as

$$y_{it} = h(\mathbf{x}_{it}, \boldsymbol{\beta}) + v_{it} + z_{it}, \quad (i = 1,\ldots, N;\ t = 1,\ldots, T) \qquad (1)$$

where y_{it} is the natural logarithm of cost for branch i at time t ($i = 1,\ldots, N$; $t = 1,\ldots, T$); $\mathbf{x}_{it}$ is a row vector of exogenous variables; h – a known measurable function and $\boldsymbol{\beta}$, – a vector of k unknown parameters define the deterministic part of the frontier and represent technology common to all branches (the translog specification is used in the empirical part); and v_{it} and z_{it} are random terms, one symmetric about zero and the other non-negative. In the case of a cost frontier, z_{it} captures the overall cost inefficiency, reflecting cost increases due to both technical and allocative inefficiency of branch i at time t. For the translog cost model, treated as the true description of technology, Kumbhakar (1997) derives the exact relationship between allocative inefficiency in the cost share equations and in the cost function, which indicates that z_{it} in (1) are not independent of the exogenous variables and the parameters in the cost function. However, the translog specification is generally viewed as an approximation to the unknown true cost

[1] In Polish efficiency studies for the banking sector, mathematical programming techniques (mainly Data Envelopment Analysis, DEA) prevail; see, e.g., Mielnik, Ławrynowicz (2002) and Pawłowska (2003a; 2003b).

[2] Rogers (1998) studied the role of nontraditional activities and their importance for measuring efficiency.

function, and the common assumption within the stochastic frontier framework is that inefficiency terms are independent of the systematic part of the cost model. Thus, this independence assumption will be maintained in our analysis.

Note that our framework is suitable for panel data, but the case of just one cross-section is easily covered as it corresponds to $T = 1$. Here we make the assumption that the inefficiency level is an individual (branch) effect, i.e. $z_{it} = z_i$ ($t = 1, \ldots, T$), as in KOS (1994b; 1997); see also Pitt, Lee (1981, Model I) and Schmidt, Sickles (1984). This assumption is motivated by our empirical example, where we use panel data corresponding to only four quarters of one year ($T = 4$). In such a short period of time, systematic changes in efficiency cannot be expected, so we use the data to improve precision of inferences on individual efficiency treated as a branch-specific characteristic. Generally, time-invariant efficiency will be measured as $r_i = \exp(-z_i)$, which is an easily interpretable quantity in (0, 1]. We also assume that z_i and v_{it} are independent of each other and v_{it} are independent across branches and time.

Using yearly data for many countries observed over longer periods, KOS (1999; 2000a; 2000b) follow an alternative strategy and assume that z_{it} is independent over both i and t (conditionally upon the parameters necessary to describe its sampling distribution); see also Pitt and Lee (1981, Model II). Osiewalski and Steel (1998) discussed numerical tools directly applicable in specifications that do not impose any panel structure. Here we use a version of the Gibbs sampler designed for random individual effects models in KOS (1997); our Gibbs sampler draws from the region (in the parameters space) where economic regularity holds.

Bayesian analysis requires specifying the Bayesian model, i.e. the joint distribution of the observables, latent variables and parameters, usually conditional on the values of the explanatory variables (assumed exogenous). According to the common statistical practice, we first formulate the sampling distribution (of the observables and latent variables given parameters) and then the prior distribution (the marginal distribution of the parameters of the sampling specification). In order to specify a parametric sampling distribution for the observables y_{it} and unobserved z_i, we assume that v_{it} is $N(0,\sigma^2)$, i.e. Normal with zero mean and constant variance σ^2, and z_i is Exponential with mean (and standard deviation) λ_i. The mean of z_i can depend on some (say, m-1) exogenous variables s_{ij} ($j = 2,\ldots, m$) explaining possible systematic differences in efficiency levels. We assume

$$\lambda_i = \prod_{j=1}^{m} \phi_j^{-s_{ij}} \qquad (2)$$

where $\phi_j > 0$ are unknown parameters and $s_{i1} = 1$. If $m > 1$, the distributions of z_i can differ for different i and thus in KOS(1997) this specification is called the Varying Efficiency Distribution (VED) model. If $m = 1$, then $\lambda_i = \phi_1^{-1}$ and all inefficiency terms constitute independent draws from the same distribution. This important special case is called the Common Efficiency Distribution (CED) model. Some non-Bayesian empirical works in the field of bank efficiency analysis used a two-step approach where the efficiency estimates obtained at the first stage were regressed (at the second stage) on additional explanatory variables; see, Cebenoyan et al. (1993), Mester (1993), Berger, Mester (1997), Berger, De Young (1997), and Kraft, Tirtiroglu (1997). While such two-step approaches can serve as very crude statistical techniques, our Bayesian VED model yields a coherent framework for both estimation and testing of influences of exogenous factors on individual efficiency.

Note that the density of all y_{it} and z_i given $\mathbf{x}_{it}$, $\mathbf{s}_i = (s_{i1},\ldots, s_{im})$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma^{-2}, \phi_1,\ldots, \phi_m)'$ is the product of NT Normal and N Exponential densities. This leads to the following Bayesian model:

$$p(\mathbf{y},\mathbf{z},\boldsymbol{\theta}|\mathbf{X},\mathbf{S}) = p(\boldsymbol{\theta})p(\mathbf{z}|\boldsymbol{\theta},\mathbf{X},\mathbf{S})p(\mathbf{y}|\mathbf{z},\boldsymbol{\theta},\mathbf{X},\mathbf{S})$$

$$\propto \; p(\boldsymbol{\theta})\prod_{i=1}^{N}\left[f_G\left(z_i \middle| 1, \prod_{j=1}^{m}\phi_j^{s_{ij}} \right) \prod_{t=1}^{T} f_N^1\left(y_{it} | h(\mathbf{x}_{it},\boldsymbol{\beta}) + z_i, \sigma^2 \right)\right] \qquad (3)$$

where $p(\boldsymbol{\theta})$ denotes the prior density, $f_N(.\,|\,a,b)$ is the (univariate) Normal density with mean a and variance b, and $f_G(.\,|\,a,b)$ is the Gamma density with mean a/b and variance a/b^2 ($a = 1$ corresponds to the Exponential distribution).

In principle, the prior distribution of $\boldsymbol{\theta}$ can be any distribution, but it is usually preferred not to introduce too much subjective information about the parameters. Therefore, we use the following prior structure:

$$p(\boldsymbol{\theta}) = p(\sigma^{-2})p(\boldsymbol{\beta})p(\boldsymbol{\phi}) \propto f_G\left(\sigma^{-2}\middle|\tfrac{1}{2}n_0, \tfrac{1}{2}c_0\right) I(\boldsymbol{\beta} \in B)\prod_{j=1}^{m} f_G(\phi_j \,|\, 1, g_j), \qquad (4)$$

which reflects the lack of prior knowledge about the frontier parameters $\boldsymbol{\beta}$, except for regularity conditions $\boldsymbol{\beta} \in B$ imposed by economic theory and represented by the indicator function $I(.)$. Alternatively, we could use a proper prior distribution on $\boldsymbol{\beta}$, possibly truncated to the region of regularity. Typically, we shall choose the prior hyperparameters $n_0 > 0$ and $c_0 > 0$ so as to represent very weak prior information on the precision of the stochastic frontier. In models without panel structure, we cannot take as the prior density for σ^{-2} the kernel of the limiting case where $c_0 = 0$ because this would result in the lack of existence of the posterior distribution; see Fernández et al. (1997). Since we treat here the inefficiency terms as time-invariant individual effects, the use of the usual Jeffreys type prior for σ^{-2} (which corresponds to the Gamma kernel with $n_0 = c_0 = 0$) is allowed. For the m parameters of the efficiency distribution we take proper, independent Exponential priors in order to avoid the pathology described by Ritter (1993) and discussed in more general terms by Fernández et al. (1997). Following KOS (1994b, 1997), we use $g_j = 1$ for $j > 1$ and take $g_1 = -\ln(r^*)$ where $r^* \in (0, 1)$ is the hyperparameter to be elicited. In the CED model ($m = 1$), r^* can be interpreted as the prior median efficiency, because it is exactly the median of the marginal prior distribution of individual efficiency $r_i = \exp(-z_i)$; see BKOS (1994). In the VED case ($m > 1$), our prior for $\phi = (\phi_1,\dots, \phi_m)'$ is quite non-informative and centered over the prior for the CED model. The prior on ϕ, a parameter which is common to all branches, induces links between the branch-specific inefficiency terms.

3. Bayesian Inference using Gibbs sampling

Since an important aspect of any empirical analysis of production is making inferences not only on the parameters describing technology, but also on individual efficiencies of observed units (here: branches), there is no need to integrate out unobserved z_i's from the joint density (3). After having observed the data, the Bayesian approach combines all the information about the unknown quantities in their posterior density $p(\mathbf{z}, \boldsymbol{\theta} \mid \mathbf{y},\mathbf{X},\mathbf{S})$ proportional to (3). As this is a non-standard and highly multivariate density, the crucial task of any applied Bayesian study is "to calculate relevant summaries of the posterior distribution, to express the posterior information in a usable form, and to serve as formal inferences if appropriate. It is in the task of summarizing that computation is typically needed." (O'Hagan 1994, p. 205). As KOS (1997) showed, a Markov Chain Monte Carlo technique known as Gibbs sampling is a particularly easy and efficient tool for simulating samples from the posterior distribution and therefore for approximating its relevant summaries.

Gibbs sampling is a technique for obtaining a sample from a joint distribution of a random vector $\boldsymbol{\alpha}$ by taking random draws from only full conditional distributions. A detailed description of the technique can be found in e.g. Casella, George (1992), and Tierney (1994). Suppose we are able to partition $\boldsymbol{\alpha}$ into $(\alpha_1', \dots, \alpha_p')'$ in such a way that sampling from each of the conditional distributions (of α_i given the remaining subvectors; $i = 1,\dots, p$) is relatively easy. Then the Gibbs sampler consists of drawing from these distributions in a cyclical way, that is, given the qth draw, $\boldsymbol{\alpha}^{(q)}$, the next draw, $\boldsymbol{\alpha}^{(q+1)}$, is obtained in the following pass through the sampler:

$\boldsymbol{\alpha}_1^{(q+1)}$ is drawn from $p(\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2= \boldsymbol{\alpha}_2^{(q)},\dots,\boldsymbol{\alpha}_p = \boldsymbol{\alpha}_p^{(q)})$,
$\boldsymbol{\alpha}_2^{(q+1)}$ is drawn from $p(\boldsymbol{\alpha}_2 \mid \boldsymbol{\alpha}_1=\boldsymbol{\alpha}_1^{(q+1)},\boldsymbol{\alpha}_3= \boldsymbol{\alpha}_3^{(q)},\dots, \boldsymbol{\alpha}_p = \boldsymbol{\alpha}_p^{(q)})$,
…
$\boldsymbol{\alpha}_p^{(q+1)}$ is drawn from $p(\boldsymbol{\alpha}_p \mid \boldsymbol{\alpha}_1=\boldsymbol{\alpha}_1^{(q+1)},\dots, \boldsymbol{\alpha}_{p-1}= \boldsymbol{\alpha}_{p-1}^{(q+1)})$.

Note that each pass consists of p steps, i.e. drawings of the p subvectors of $\boldsymbol{\alpha}$. The starting point, $\boldsymbol{\alpha}^{(0)}$, is arbitrary. Under certain general conditions (irreducibility and aperiodicity as described in e.g. Tierney (1994), the distribution of $\boldsymbol{\alpha}^{(q)}$ converges to the joint distribution, $p(\boldsymbol{\alpha})$, as q tends to infinity. Thus, in an asymptotic sense, we draw a sample directly from the joint distribution. In practical applications we have to discard a (large) number of passes before convergence to joint distribution $p(\boldsymbol{\alpha})$ is reached.

In order to efficiently use Gibbs sampling to make posterior inferences on both the parameters and branch efficiencies, we have to consider the joint posterior density of $\mathbf{z}$ and $\boldsymbol{\theta}$, $p(\mathbf{z}, \boldsymbol{\theta} \mid \mathbf{y},\mathbf{X},\mathbf{S})$ where z is the $N \times 1$ vector of all the z_is. Note that the dimension is then $N + k + m + 1$, greater than the number of observed units. Despite this high dimensionality, the steps involved in the Gibbs sampler are very easy to implement.

Given $\mathbf{z}$, the frontier parameters $(\boldsymbol{\beta},\sigma^{-2})$ are independent of ϕ and can be treated as the parameters of the (linear or nonlinear) Normal regression model in (3). Thus, we obtain the following full conditionals for σ^{-2} and $\boldsymbol{\beta}$:

$$p(\sigma^{-2} \mid \mathbf{y},\mathbf{X},\mathbf{S},\mathbf{z},\boldsymbol{\beta},\boldsymbol{\phi}) = p(\sigma^{-2} \mid \mathbf{y},\mathbf{X},\mathbf{z},\boldsymbol{\beta}) = f_G\left(\sigma^{-2} \middle| \frac{n_0 + T\cdot N}{2}, \frac{1}{2}\left\{c_0 + \sum_{i,t}\left(y_{it} - z_i - h(\mathbf{x}_{it},\boldsymbol{\beta})\right)^2\right\}\right) \tag{5}$$

$$p(\boldsymbol{\beta} \mid \mathbf{y},\mathbf{X},\mathbf{S},\mathbf{z},\sigma^{-2},\boldsymbol{\phi}) = p(\boldsymbol{\beta} \mid \mathbf{y},\mathbf{X},\mathbf{z},\sigma^{-2}) \propto I(\boldsymbol{\beta} \in B)\exp\left\{-\tfrac{1}{2}\sigma^{-2}\sum_{i,t}\left(y_{it} - z_i - h(\mathbf{x}_{it},\boldsymbol{\beta})\right)^2\right\}. \tag{6}$$

The full conditional posterior densities of ϕ_j ($j = 1, \dots, m$) have the general form:

$$p(\phi_j \mid \mathbf{y},\mathbf{X},\mathbf{S},\mathbf{z},\boldsymbol{\beta},\sigma^{-2},\boldsymbol{\phi}_{(-j)}) = p(\phi_j \mid \mathbf{S},\boldsymbol{\phi}_{(-j)}) \propto \exp\left(-\phi_1\sum_i z_i D_{i1}\right)\times f_G\left(\phi_j \middle| 1 + \sum_i s_{ij}, g_j\right) \tag{7}$$

where

$$D_{ir} = \prod_{j\neq r}^{m} \phi_j^{s_{ij}} \tag{8}$$

for $r = 1, \dots, m$ ($D_{i1} = 1$ when $m = 1$) and $\boldsymbol{\phi}_{(-j)}$ denotes $\boldsymbol{\phi}$ without its jth element. Since $s_{i1} = 1$, the conditional of ϕ_1 is Gamma with parameters $1 + N$ and $g_1 + z_1 D_{11} + \dots + z_N D_{N1}$.

Conditionally on the parameters and the data, the vector of unobserved inefficiency terms $\mathbf{z} = (z_1 \dots z_N)'$ has a truncated Normal distribution with density which is the product of N independent truncated Normal

$$p(\mathbf{z}|\mathbf{y},\mathbf{X},\mathbf{S},\boldsymbol{\theta}) \propto \prod_{i=1}^{N} f_N^1\left(z_i \middle| \bar{y}_i - \bar{\mathbf{x}}_i\boldsymbol{\beta} - T^{-1}\sigma^2\prod_{j=1}^{m}\phi_j^{s_{ij}}, T^{-1}\sigma^2\right) I(z_i \geq 0) \tag{9}$$

densities; see KOS(1997). In (9) $\bar{y}_i$ and $\bar{x}_i$ are simple averages of y_{it} and $\mathbf{x}_{it}$ over time. From (9) we can easily draw z_is given the data and the parameters. These draws are immediately transformed into efficiency indicators $r_i = \exp(-z_i)$. Thus, this N-dimensional step of each pass through our Gibbs sampler is quite simple.

Depending on the form of the frontier and on the values of s_{ij} for $j > 1$, the full conditionals for $\boldsymbol{\beta}$ and for ϕ_j ($j = 2, \ldots, m$) can be quite easy or very difficult to draw from. Drawing from nonstandard conditional densities within the Gibbs sampler requires special techniques, like rejection methods or the Metropolis-Hastings algorithm (see e.g. Tierney 1994 or O'Hagan 1994). KOS (1994a; 1994b) used variants of the Metropolis-Hastings technique in the cases of a non-linear frontier or continuous s_{ij}, respectively. These hybrid procedures imply a substantial added complexity in simulations from the posterior distribution and require additional input from the user. Therefore, following Osiewalski and Steel (1998) we stress two special cases where considerable simplifications are possible:

(i) linearity of the frontier,

(ii) 0–1 dummies for s_{ij} ($j = 2, \ldots, m$).

If $h(\mathbf{x}_{ti}, \boldsymbol{\beta}) = \mathbf{x}_{ti}\boldsymbol{\beta}$ then (6) is a k-variate Normal density, possibly truncated due to regularity conditions. That is, we have

$$p\left(\boldsymbol{\beta} \mid \mathbf{y}, \mathbf{X}, \mathbf{S}, \mathbf{z}, \sigma^{-2}, \boldsymbol{\phi}\right) \propto I(\boldsymbol{\beta} \in B) f_N^k\left(\boldsymbol{\beta} \mid \hat{\boldsymbol{\beta}}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}\right) \quad (10)$$

where

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} - \mathbf{z} \otimes \boldsymbol{\iota}_T), \quad (11)$$

$\boldsymbol{\iota}_T$ is a vector of ones, $\mathbf{z}\otimes\boldsymbol{\iota}_T$ denotes the Kronecker product of $\mathbf{z}$ and $\boldsymbol{\iota}_T$, and $\mathbf{y}$ and $\mathbf{X}$ denote a $NT \times 1$ vector of y_{it}s and a $NT \times k$ matrix with $\mathbf{x}_{it}$s as rows, respectively. Cobb-Douglas or translog frontiers serve as examples of linearity in $\boldsymbol{\beta}$; see Koop et al. (1995) and KOS (1997; 1999; 2000a; 2000b).

The dichotomous character of the variables explaining efficiency differences ($s_{i2}, \ldots, s_{im}$) greatly simplifies (7), which simply becomes a Gamma density:

$$p\left(\phi_h \mid \mathbf{S}, \mathbf{z}, \boldsymbol{\phi}_{(-h)}\right) = f_G\left(\phi_h \left| 1 + \sum_{i=1}^{N} s_{ih}, g_h + \sum_{i=1}^{N} s_{ih} z_i D_{ih}\right.\right); \quad (12)$$

see KOS (1997). From the purely numerical perspective, it pays to dichotomize these original variables in $\mathbf{s}_i$ which are not 0–1 dummies.

The above discussion confirms that the Bayesian stochastic frontier cost model, considered in this paper, can be analyzed using Gibbs sampling. That is, even though the marginal posteriors of $\boldsymbol{\theta}$ and z_i are unwieldy, the conditionals for a suitable partition of the set of unknown quantities are much easier to work with. By taking a long enough sequence of successive draws from the conditional posterior densities, each conditional on previous draws from the other conditional densities, we can create a sample that can be treated as coming from the joint posterior distribution. The posterior expectation of any arbitrary function of interest, $g(\boldsymbol{\theta}, \mathbf{z}; \mathbf{y}, \mathbf{X}, \mathbf{S})$, can be approximated by its sample mean, g^*, based on M passes (after convergence has been assured):

$$E\left[g(\boldsymbol{\theta}, \mathbf{z}; \mathbf{y}, \mathbf{X}, \mathbf{S}) \middle| \mathbf{y}, \mathbf{X}, \mathbf{S}\right] \approx g^*(\mathbf{y}, \mathbf{X}, \mathbf{S}) = \frac{1}{M} \sum_{l=1}^{M} g\left(\boldsymbol{\theta}^{(l)}, \mathbf{z}^{(l)}; \mathbf{y}, \mathbf{X}, \mathbf{S}\right).$$

4. Modelling Variable Costs of Bank Branches

We illustrate the Bayesian stochastic frontier analysis using the data from $N = 58$ branches of one of Polish commercial banks observed over $T = 4$ quarters of one year. Our short-run translog cost model takes the form:

$$\begin{aligned} \ln VC_{it} = {} & \beta_0 + \beta_1 \ln W_{it,D} + \beta_2 \ln W_{it,L} + \beta_3 \ln Q_{it} + \beta_4 \ln K_i + \beta_5 \ln W_{it,D} \ln W_{it,L} \\ & + \beta_6 \ln W_{it,D} \ln Q_{it} + \beta_7 \ln W_{it,D} \ln K_i + \beta_8 \ln W_{it,L} \ln Q_{it} + \beta_9 \ln W_{it,L} \ln K_i \\ & + \beta_{10} \ln Q_{it} \ln K_i + \beta_{11} (\ln W_{it,D})^2 + \beta_{12} (\ln W_{it,L})^2 + \beta_{13} (\ln Q_{it})^2 \quad (14) \\ & + \beta_{14} (\ln K_i)^2 + v_{it} + z_i \end{aligned}$$

where the following notation is adopted:

VC = cost of labour (personnel expenses) + cost of financial capital (interest expenses) + cost of computers, software and other goods and services purchased from outside suppliers,

W_L = price of labour = (personnel expenses)/(number of full-time equivalent employees),

W_D = price of deposits and other borrowed money = (interest expense)/(volume),

K = office space (in square meters),

Q = aggregate volume of different loans + the excess of deposits over loans (if positive).

In our VED specification for inefficiency term z_i, we use three dummies to model its mean λ_i:

$s_{i2} = 1$ if branch i had more deposits than loans ($s_{i2} = 0$ otherwise),

$s_{i3} = 1$ if volume of loans was greater than PLN 100 million ($s_{i3} = 0$ otherwise),

$s_{i4} = 1$ for branches with subbranches ($s_{i4} = 0$ otherwise);

thus, $m = 4$ and $s_{i1} \equiv 1$. Our conjecture is that fewer deposits than loans means higher costs because of charge for "external" refinancing (thus, $s_{i2} = 1$ should correspond to higher efficiency, $\phi_2 > 1$); the larger branch and the more complicated its structure, the lower efficiency ($\phi_3 < 1$, $\phi_4 < 1$). The model has 20 parameters, including σ^{-2}, ϕ_1, ϕ_2, ϕ_3 and ϕ_4.

In the specification given above, we follow the microeconomic analysis of Sealey and Lindley (1977) who view the bank as using labour, physical capital, and financial capital (mainly deposits) to produce earning assets. Thus, we use deposits and other borrowed money as inputs (representing financial capital used), and income generating money as the aggregate product of a bank branch. This approach (often called the intermediation approach) has been adopted in many empirical studies, using econometric as well as mathematical programming tools; see Akhaiven et al. (1997), Altunbas et al. (2000), Berger et al. (1997), Cebenoyan et al. (1993), English et al. (1993), Grabowski et al. (1993), Hassan et al. (1990), Hughes, Mester (1993), Kaparakis et al. (1994), Mester (1987; 1993), Muldur, Sassenou (1993), Noulas et al. (1990), Zardokoohi, Kolari (1994).

As a consequence of the approach we follow, the variable cost includes both interest and operating costs. Our aggregate product Q comprises loans to individuals, commercial and industrial loans, and the excess of deposits over loans (if positive, i.e. if $s_{i2} = 1$). The latter component reflects the fact that branches operate within the bank and their excessive deposits can be used by those branches which lack funds for loans ($s_{i2} = 0$). In fact, all observed branches tended to specialize either in the acquisition of financial capital from depositors or in lending funds. Branches from the first group (depository branches, $s_{i2} = 1$) provided extra funds, which were used by branches from the other group. These funds were provided at a constant price (related to prices on the interbank market), fixed by the bank and only used to correct the calculation of the operating profit of a branch. Thus, for a depository branch, the volume of its excess funds can be treated as a product because it increased the calculated profit of that branch. On the other hand, this money was used as input by the branches that lacked funds for loans; its price was constant over branches and used to correct downwards the calculation of the operating profit of branches specializing in lending funds.

Our measure of variable cost includes cost of computers, software and other goods and services purchased from outside suppliers but their prices do not appear as explanatory variables in our specification. This is a consequence of the fact that these prices can be treated as constant (over the whole year and all the branches) as main purchases were decided on the level of the bank which chose a supplier (of e.g. hardware or software) once during several months. Thus, the effect of these prices on the variable cost is taken by five of the parameters β_i ($i = 0, 1, 2, 3, 4$). However, the elasticities of VC with respect to W_D, W_L, Q and K calculated from the full translog model remain unaffected by unobservability of constant prices and are the same as calculated from (14). Moreover, homogeneity with respect to all prices is automatically fulfilled by (14).

Other regularity conditions which should be imposed on our specification include monotonicity (with respect to Q and all prices) and concavity (in all prices). Under any continuous prior distribution on the parameter space, monotonicity with probability one is equivalent to positivity of elasticities of VC with respect to Q, W_D and W_L plus the condition that the sum of elasticities with respect to W_D and W_L is less than one:

$$\begin{aligned}
&\eta(VC \mid Q) = \beta_3 + \beta_6 \overline{\ln W_D} + \beta_8 \overline{\ln W_L} + \beta_{10} \overline{\ln K} + 2\beta_{13} \overline{\ln Q} > 0,\\
&\eta(VC \mid W_D) = \beta_1 + \beta_5 \overline{\ln W_L} + \beta_6 \overline{\ln Q} + \beta_7 \overline{\ln K} + 2\beta_{11} \overline{\ln W_D} > 0,\\
&\eta(VC \mid W_L) = \beta_2 + \beta_5 \overline{\ln W_D} + \beta_8 \overline{\ln Q} + \beta_9 \overline{\ln K} + 2\beta_{12} \overline{\ln W_L} > 0,\\
&\eta(VC \mid W_D) + \eta(VC \mid W_L) < 1.
\end{aligned} \tag{15}$$

The latter condition assures that VC is increasing in unobservable (constant) prices. Remind that the elasticities with respect to prices are equal to optimal shares (of production factors) in variable cost. We impose the monotonicity and concavity conditions on an "average" branch, that is a hypothetical branch with average (over time and branches) values of logs of K, Q, W_D and W_L. Concavity in input prices is equivalent to negative semi-definiteness of the matrix of second order derivatives of VC with respect to all three prices. Since this matrix is singular, it is negative semi-definite iff all three first order principal minors are non-positive and all three second order principal minors are non-negative; see e.g. Simon, Blume (1994). Under any continuous prior distribution on the parameter space, prior and posterior probabilities of equalities are zero. Thus, concavity in prices is assured with (both prior and posterior) probability one iff the first two leading principal minors change sign, i.e. if

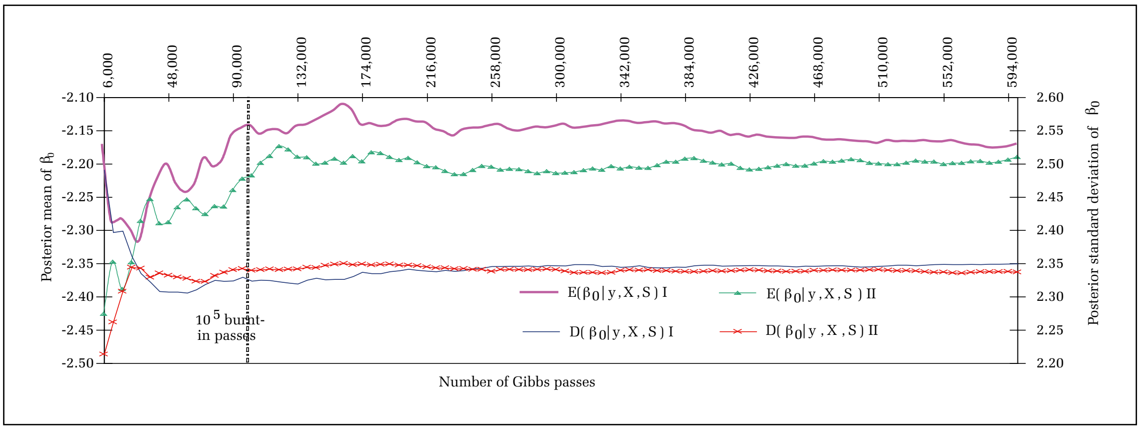
$$\begin{aligned}
&2\beta_{11} + \eta(VC \mid W_D)\cdot(\eta(VC \mid W_D) - 1) < 0,\\
&(2\beta_{11} + \eta(VC \mid W_D)\cdot(\eta(VC \mid W_D) - 1))\cdot(2\beta_{12} + \eta(VC \mid W_L)\cdot(\eta(VC \mid W_L) - 1))\\
&-(\beta_5 + \eta(VC \mid W_D)\cdot\eta(VC \mid W_L))^2 > 0.
\end{aligned} \tag{16}$$

It is easy to prove that, given(16) and positivity of elasticities, all principal minors (not only the leading principal minors) have correct signs.

We impose regularity at a particular point in the space of explanatory variables (at the point where our translog specification should best approximate the unknown cost function). Although economic regularity could be imposed at many points, that would lead to reducing flexibility of the translog approximation and a serious increase in computational burden.

In fact, the first three restrictions in (15) are not binding for our data set. Moreover, the elasticities with respect to Q and W_D are clearly positive for all 58 branches and the elasticity with respect to W_L always has a positive posterior mean. The fourth restriction is binding, as will be shown below; see Table 2 and Figure 4. Given the monotonicity alone, the posterior probability of concavity is only 0.11, so the two inequalities presented in (16) are obviously binding. Salvanes and Tjøtta (1998) illustrate importance of the concavity restrictions for the interpretation of the results of cost function estimation.

Figure 1. *Gibbs estimates of the posterior mean and standard deviation of β_0 as functions of the number of passes (for two different runs)*



Source: Authors' calculations.

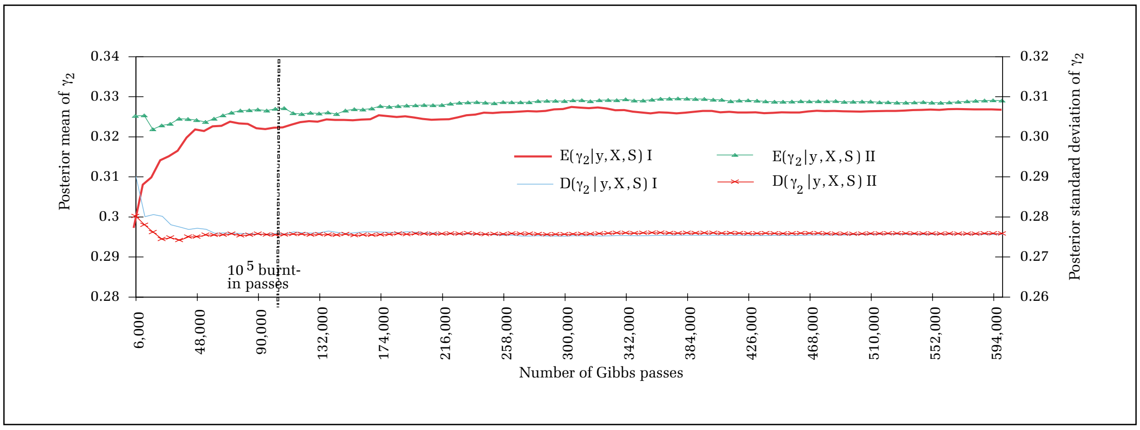
In our short-run model we represent physical capital (treated as fixed input) by the office space used by the branches. We have also estimated specifications with K defined as the book value of buildings and offices, obtaining very similar results.

As regards the prior hyperparameters, we set $r^* = 0.7$ which would be the prior median of efficiency in the CED specification ($m = 1$, no systematic differences in cost efficiency). Thus, in our VED specification with $m = 4$ we assume prior median efficiency even lower than 0.7 (about 0.53, as obtained by Marzec (2000) for average values of s_{i2}, s_{i3}, s_{i4}). Other values of r^* from the interval [0.5, 0.9] have no new consequences for our inference on technology but show some (although small) influence on the efficiency analysis discussed in the next section. For precision of the Normal error term, we take $n_0 = c_0 = 10^{-6}$ which leads to a very diffuse Gamma prior distribution (with mean 1 and variance 2×10^6) reflecting little prior knowledge about this parameter. Assuming the improper prior corresponding to $n_0 = c_0 = 0$ leads to the same posterior results.

The Gibbs sampler presented in the previous section requires starting values for $\boldsymbol{\beta}$ and $\mathbf{z}$. We tried different vectors $\boldsymbol{\beta}^{(0)}$ and $\mathbf{z}^{(0)}$, receiving virtually the same results after about 100,000 passes. In particular, in some runs we used the same $z_i^{(0)}$ (e.g., 0.3) for all i and we calculated $\boldsymbol{\beta}^{(0)}$ from the OLS formula (11). Convergence to the posterior distribution is illustrated in Figures 1 and 2, which show changes (in two Gibbs runs starting from very different initial values) of the Monte Carlo estimates of posterior means and standard deviations for two parameters with particularly slow convergence. Note that the estimates of standard deviations are closer to each other than the estimates of posterior means. Even in the latter case, however, the differences in the final estimates are as small as 1% of the corresponding posterior standard deviation.

The posterior results for our 78-dimensional vector of unknown parameters and inefficiency terms were obtained using one long run of 500,000 Gibbs passes,

Figure 2. *Gibbs estimates of the posterior mean and standard deviation of γ_2 as functions of the number of passes (for two different runs)*



Source: Authors' calculations.

Table 1. *Posterior means and standard deviations of the parameters of model (14) (VED with m = 4; r* = 0.7)*

Parameter	Variable	E(* | data)	D(* | data)
β_0	Constant	-2.178	2.355
β_1	$\ln W_D$	1.263	0.425
β_2	$\ln W_L$	0.383	0.345
β_3	Ln_Q	0.619	0.225
β_4	Ln_K	-0.388	0.178
β_5	$\ln W_D \ln W_L$	-0.040	0.039
β_6	$\ln W_D \ln Q$	-0.014	0.025
β_7	$\ln W_D \ln K$	0.048	0.023
β_8	lnWL lnQ	-0.032	0.015
β_9	lnWL lnK	0.005	0.011
β_{10}	lnQ lnK	-0.012	0.011
β_{11}	$(\ln WD)^2$	-0.048	0.035
β_{12}	$(\ln WL)^2$	0.029	0.021
β_{13}	$(\ln Q)^2$	0.017	0.008
β_{14}	$(\ln K)^2$	0.035	0.009
ϕ_1	Constant (s_{i1}=1)	11.522	3.007
ϕ_2	si_2	1.440	0.397
ϕ_3	si_3	0.820	0.256
ϕ_4	si_4	0.949	0.393
σ^2	-	2.83×10^{-4}	0.38×10^{-4}

Source: Authors' calculations.

after discarding 100,000 initial draws. Tables 1, 2 and 3 present the posterior means and standard deviations of the parameters of the frontier cost function, the elasticities for the "average" branch and the elasticities for all branches (ordered by decreasing production), respectively. The individual elasticities in Table 3 are estimated assuming time averages for explanatory variables (expressed in logs). As regards factor prices, the interest rate on deposits (i.e. the price of financial capital) exerts the strongest influence on variable cost; the role of the price of labour is much smaller. Note that we can write the sum of elasticities with respect to those factor prices that are constant over branches as one minus the sum of elasticities with respect to W_D and W_L. Figures 3 and 4 show the (very sharp) marginal posterior densities of the variable cost elasticities for the "average" branch.[3]

Table 3 clearly shows that elasticities vary a lot over branches, making the Cobb–Douglas specification completely inadequate. Also the functional form suggested by Nerlove (1963) and used by Christensen, Greene (1976), BKOS (1994) and Osiewalski, Marzec (1998), which is based on the Cobb-Douglas specification but permits returns to scale to vary with Q, is not supported by the data. Let $\boldsymbol{\beta}^* = (\beta_5\ \beta_6\ \beta_7\ \beta_8\ \beta_9\ \beta_{10}\ \beta_{11}\ \beta_{12}\ \beta_{14})'$; since the marginal posterior of $\boldsymbol{\beta}^*$ is approximately Normal with mean $E(\boldsymbol{\beta}^* | \mathbf{y}, \mathbf{X}, \mathbf{S})$ and covariance matrix $V(\boldsymbol{\beta}^* \mid \mathbf{y}, \mathbf{X}, \mathbf{S})$, the posterior of $\tau(\boldsymbol{\beta}^*; \mathbf{y}, \mathbf{X}, \mathbf{S}) = [\boldsymbol{\beta}^* - E(\boldsymbol{\beta}^* \mid \mathbf{y}, \mathbf{X}, \mathbf{S})]' V^{-1}(\boldsymbol{\beta}^* \mid \mathbf{y}, \mathbf{X}, \mathbf{S}) [\boldsymbol{\beta}^* - E(\boldsymbol{\beta}^* \mid \mathbf{y}, \mathbf{X}, \mathbf{S})]$ is close to the chi-square distribution with 9 degrees of freedom. The value $\tau(\mathbf{0}; \mathbf{y}, \mathbf{X}, \mathbf{S})$, corresponding to the simpler functional form, is equal to 256,381 and lies very far in the tail of the posterior density of $\tau(\boldsymbol{\beta}^*; \mathbf{y}, \mathbf{X}, \mathbf{S})$.

From Table 1 we see that the elasticity of VC with respect to Q, $\eta(VC \mid Q)$ increases significantly with Q ($\beta_{13} > 0$) but decreases with W_L ($\beta_8 < 0$); the elasticity of VC with respect to W_D increases with K ($\beta_7 > 0$); the elasticity of VC with respect to W_L decreases with Q ($\beta_8 < 0$); the elasticity of VC with respect to K increases significantly with K ($\beta_{14} > 0$) and W_D ($\beta_7 > 0$). Figure 5 presents the posterior mean of $\eta(VC \mid Q)$ as a function

[3] In fact, we performed several different very long Gibbs runs in order to check numerical stability of our results. The striking similarity of all posterior characteristics in all runs illustrates convergence of the Gibbs sampler.

Table 2. *Posterior means and standard deviations of elasticities for the "average" branch (VED, m = 4; r* = 0.7)*

	$\eta(VC\|W_D)$	$\eta(VC\|W_L)$	$\eta(VC\|Q)$	$\eta(VC\|K)$	$1-\eta(VC\|W_L)-\eta(VC\|W_D)$
Means	0.797	0.186	0.863	0.036	0.017
Standard deviations	(0.013)	(0.010)	(0.012)	(0.009)	(0.011)

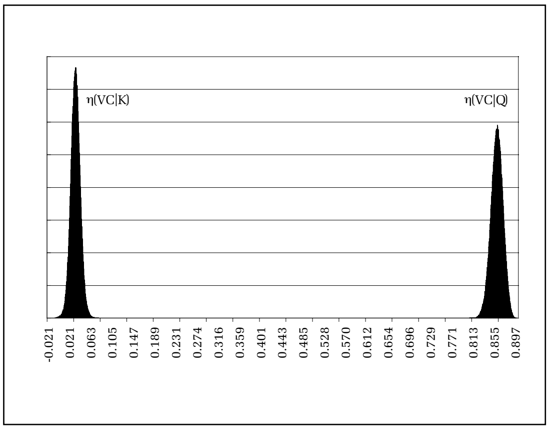
Source: Authors' calculations.

Table 3. *Posterior means and standard deviations (in parentheses) of elasticities for all branches*

i	$\eta(VC\|W_D)_i$		$\eta(VC\|W_L)_i$		$1-[\eta(VC\|W_D)_i+\eta(VC\|(W_L)_i]$		$\eta(VC\|Q)_i$		$\eta(VC\|K)_i$	
1	0.793	(0.051)	0.126	(0.029)	0.081	(0.046)	0.906	(0.034)	0.091	(0.027)
2	0.814	(0.035)	0.146	(0.025)	0.040	(0.036)	0.914	(0.033)	0.035	(0.022)
3	0.720	(0.059)	0.134	(0.033)	0.146	(0.056)	0.925	(0.043)	-0.067	(0.040)
4	0.823	(0.028)	0.160	(0.018)	0.017	(0.026)	0.889	(0.024)	0.080	(0.019)
5	0.800	(0.031)	0.155	(0.018)	0.045	(0.028)	0.890	(0.025)	0.068	(0.019)
6	0.805	(0.028)	0.164	(0.017)	0.032	(0.026)	0.881	(0.022)	0.084	(0.018)
7	0.843	(0.038)	0.189	(0.029)	-0.031	(0.035)	0.898	(0.027)	0.002	(0.020)
8	0.820	(0.024)	0.166	(0.016)	0.014	(0.021)	0.885	(0.021)	0.068	(0.016)
9	0.834	(0.035)	0.156	(0.020)	0.010	(0.031)	0.872	(0.020)	0.137	(0.027)
10	0.847	(0.030)	0.168	(0.017)	-0.014	(0.026)	0.869	(0.018)	0.127	(0.024)
11	0.801	(0.021)	0.167	(0.015)	0.031	(0.020)	0.889	(0.022)	0.023	(0.016)
12	0.775	(0.052)	0.145	(0.029)	0.080	(0.045)	0.863	(0.024)	0.131	(0.027)
13	0.832	(0.022)	0.177	(0.014)	-0.010	(0.019)	0.870	(0.016)	0.087	(0.016)
14	0.701	(0.059)	0.148	(0.029)	0.151	(0.050)	0.894	(0.033)	-0.045	(0.033)
15	0.837	(0.021)	0.175	(0.014)	-0.012	(0.018)	0.873	(0.016)	0.080	(0.015)
16	0.824	(0.025)	0.170	(0.015)	0.006	(0.022)	0.865	(0.016)	0.108	(0.019)
17	0.794	(0.025)	0.171	(0.015)	0.035	(0.022)	0.862	(0.015)	0.085	(0.015)
18	0.902	(0.045)	0.194	(0.023)	-0.096	(0.039)	0.857	(0.018)	0.138	(0.028)
19	0.767	(0.026)	0.170	(0.015)	0.064	(0.022)	0.876	(0.020)	0.015	(0.016)
20	0.883	(0.038)	0.194	(0.021)	-0.076	(0.032)	0.869	(0.017)	0.078	(0.017)
21	0.784	(0.022)	0.175	(0.013)	0.041	(0.019)	0.863	(0.015)	0.062	(0.013)
22	0.812	(0.015)	0.180	(0.011)	0.008	(0.012)	0.870	(0.014)	0.046	(0.010)
23	0.751	(0.035)	0.183	(0.023)	0.066	(0.035)	0.893	(0.028)	-0.087	(0.035)
24	0.826	(0.024)	0.174	(0.016)	0.000	(0.023)	0.856	(0.013)	0.113	(0.019)
25	0.777	(0.019)	0.193	(0.016)	0.030	(0.018)	0.874	(0.018)	-0.020	(0.019)
26	0.804	(0.022)	0.187	(0.013)	0.009	(0.022)	0.848	(0.012)	0.100	(0.016)
27	0.854	(0.027)	0.197	(0.016)	-0.050	(0.023)	0.862	(0.013)	0.067	(0.012)
28	0.713	(0.041)	0.180	(0.022)	0.107	(0.037)	0.878	(0.026)	-0.065	(0.032)
29	0.769	(0.026)	0.198	(0.019)	0.034	(0.025)	0.877	(0.021)	-0.054	(0.026)
30	0.768	(0.022)	0.180	(0.013)	0.052	(0.019)	0.863	(0.015)	0.023	(0.013)
31	0.822	(0.016)	0.186	(0.012)	-0.008	(0.016)	0.856	(0.010)	0.078	(0.012)
32	0.859	(0.032)	0.194	(0.018)	-0.053	(0.031)	0.842	(0.014)	0.140	(0.026)
33	0.804	(0.028)	0.200	(0.020)	-0.004	(0.026)	0.878	(0.020)	-0.040	(0.022)
34	0.786	(0.021)	0.200	(0.016)	0.014	(0.019)	0.870	(0.017)	-0.025	(0.019)
35	0.778	(0.017)	0.190	(0.012)	0.032	(0.015)	0.866	(0.014)	0.000	(0.014)
36	0.781	(0.018)	0.186	(0.013)	0.034	(0.016)	0.869	(0.015)	-0.004	(0.015)
37	0.830	(0.020)	0.201	(0.012)	-0.031	(0.019)	0.850	(0.009)	0.063	(0.009)
38	0.757	(0.027)	0.184	(0.017)	0.059	(0.024)	0.858	(0.014)	0.004	(0.015)
39	0.835	(0.021)	0.201	(0.013)	-0.035	(0.020)	0.850	(0.009)	0.056	(0.008)
40	0.873	(0.035)	0.204	(0.019)	-0.077	(0.033)	0.839	(0.014)	0.115	(0.021)
41	0.766	(0.027)	0.190	(0.018)	0.045	(0.027)	0.839	(0.012)	0.063	(0.011)
42	0.744	(0.033)	0.197	(0.019)	0.058	(0.030)	0.868	(0.020)	-0.070	(0.030)
43	0.805	(0.018)	0.209	(0.012)	-0.014	(0.018)	0.844	(0.009)	0.040	(0.006)
44	0.795	(0.020)	0.194	(0.015)	0.011	(0.022)	0.840	(0.009)	0.068	(0.009)
45	0.771	(0.025)	0.213	(0.017)	0.016	(0.023)	0.855	(0.014)	-0.033	(0.020)
46	0.747	(0.031)	0.188	(0.019)	0.066	(0.028)	0.851	(0.014)	0.001	(0.016)
47	0.805	(0.018)	0.206	(0.012)	-0.011	(0.019)	0.845	(0.009)	0.031	(0.007)
48	0.842	(0.025)	0.205	(0.016)	-0.047	(0.027)	0.837	(0.011)	0.082	(0.013)
49	0.750	(0.034)	0.195	(0.020)	0.056	(0.032)	0.865	(0.018)	-0.066	(0.029)
50	0.750	(0.029)	0.197	(0.018)	0.053	(0.027)	0.843	(0.012)	0.009	(0.013)
51	0.762	(0.034)	0.216	(0.021)	0.022	(0.032)	0.857	(0.018)	-0.073	(0.029)
52	0.789	(0.021)	0.207	(0.015)	0.004	(0.022)	0.843	(0.010)	0.011	(0.010)
53	0.723	(0.041)	0.194	(0.023)	0.083	(0.037)	0.852	(0.018)	-0.050	(0.027)
54	0.826	(0.032)	0.226	(0.018)	-0.052	(0.031)	0.827	(0.013)	0.054	(0.009)
55	0.742	(0.035)	0.205	(0.021)	0.053	(0.033)	0.850	(0.016)	-0.057	(0.026)
56	0.824	(0.034)	0.222	(0.022)	-0.045	(0.039)	0.810	(0.018)	0.108	(0.019)
57	0.766	(0.031)	0.220	(0.021)	0.014	(0.034)	0.821	(0.014)	0.012	(0.012)
58	0.820	(0.050)	0.254	(0.030)	-0.073	(0.054)	0.792	(0.026)	0.056	(0.017)
Averages of posterior means and standard deviations										
	0.797	(0.030)	0.186	(0.018)	0.983	(0.028)	0.863	(0.018)	0.036	(0.019)

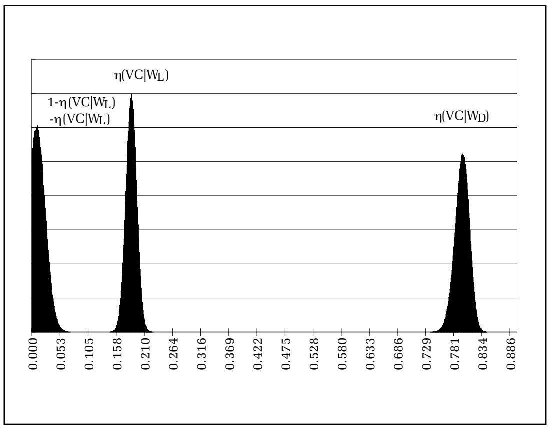
Source: Authors' calculations.

Figure 3. *Posterior distributions of elasticities of VC with respect to K and Q for the "average" branch*



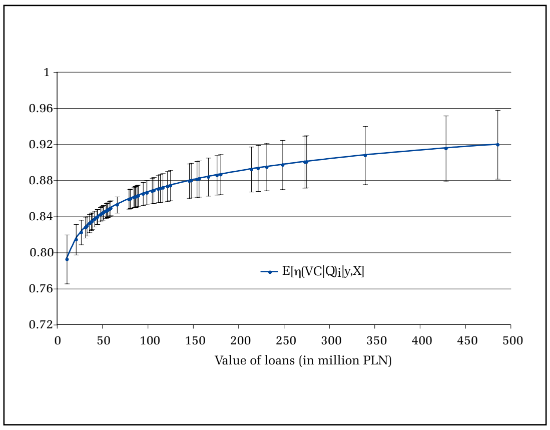
Source: Authors' calculations.

Figure 4. *Posterior distributions of elasticities of VC with respect to input prices for the "average" branch*



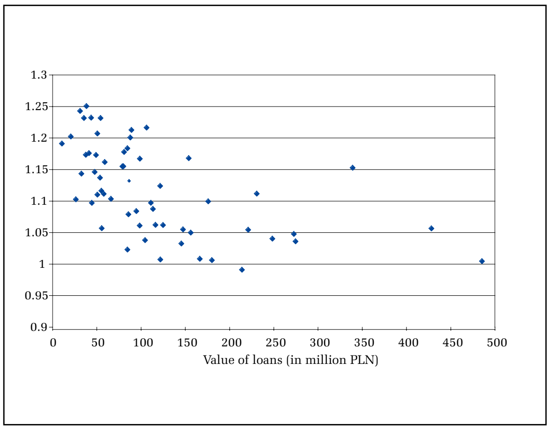
Source: Authors' calculations.

Figure 5. *Posterior means (and standard deviations) of the elasticity of VC with respect to Q as a function of Q alone (for average values of logs of other variables)*



Source: Authors' calculations.

Figure 6. *Returns to scale estimates for all branches (plotted against the output level)*

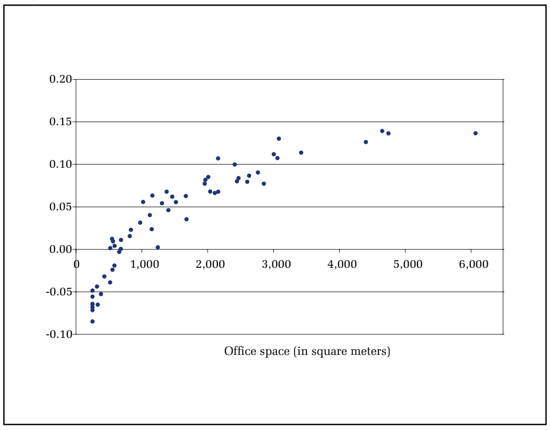


Source: Authors' calculations.

of *Q* alone, keeping the other three arguments fixed at the 'average' branch levels; we also present the posterior standard deviations at particular points. The variability of η(*VC* | *Q*) is closely related to the variability of returns to scale, defined as $RTS=(1-\partial \ln VC/\partial \ln K)/(\partial \ln VC/\partial \ln Q)$; *RTS*, evaluated for all branches at the posterior means of β_i, is presented in Figure 6. For most branches, our estimates of *RTS* are greater than one. Since all branches made profits in the observed year, most of them could have been more profitable just by increasing scale of their activities. This holds especially for small branches.

The positive elasticity with respect to the fixed factor, observed in Figure 7 for most branches (especially the ones with very large office space), suggests that these branches are far from long-run cost minimisation. This means that short-run cost efficiency, calculated on the

Figure 7. *Posterior means of fixed input elasticities for all branches (plotted against the fixed input level)*



Source: Authors' calculations.

basis of variable costs and presented in the next section, is higher than long-run efficiency of branches with too much office space.

5. Inference on Short-Run Cost Efficiency

We assumed that *a priori* there is more than 50% chance that variable cost efficiency of any given branch is below $r^* = 0.7$. Our data set points at much higher efficiency and leads to the average posterior mean of r_i equal to 0.919 with 0.017 as the average posterior standard deviation. However, the individual posterior means (and standard deviations) are quite spread, ranging from 0.768 (0.016) for branch No. 58 to 0.996 (0.004) for branch No. 56; see Table 4. As branch No. 56 has significantly positive elasticity of variable cost with respect to fixed input, its long-run cost efficiency is probably much lower. Therefore, we can treat branch No. 45 as the leading branch, with almost as high short-run cost efficiency as branch No. 56, but with the negative posterior mean of the elasticity with respect to *K*. Other branches with very high short-run efficiencies and negative elasticities with respect to *K* are branches No. 55, 33, 42, 34, 36 and 49. They all belong to the group of smaller branches.

Table 4. *Posterior means (and standard deviations) for λ_i and individual efficiency levels (VED with m = 4; r* = 0.7)*

i	s_{i2}	s_{i3}	s_{i4}	λ_i	r_i	i	s_{i2}	s_{i3}	s_{i4}	λ_i	r_i
1	1	1	1	0.105 (0.053)	0.895 (0.036)	30	1	0	0	0.066 (0.016)	0.857 (0.014)
2	0	1	0	0.119 (0.034)	0.898 (0.030)	31	0	0	0	0.093 (0.025)	0.956 (0.015)
3	0	1	0	0.119 (0.034)	0.871 (0.052)	32	0	0	0	0.093 (0.025)	0.909 (0.017)
4	0	1	1	0.144 (0.067)	0.899 (0.021)	33	1	0	0	0.066 (0.016)	0.974 (0.013)
5	0	1	0	0.119 (0.034)	0.891 (0.021)	34	0	0	0	0.093 (0.025)	0.966 (0.013)
6	0	1	0	0.119 (0.034)	0.926 (0.020)	35	1	0	0	0.066 (0.016)	0.967 (0.014)
7	0	1	0	0.119 (0.034)	0.984 (0.014)	36	0	0	0	0.093 (0.025)	0.960 (0.014)
8	0	1	0	0.119 (0.034)	0.845 (0.017)	37	1	0	0	0.066 (0.016)	0.837 (0.013)
9	0	1	1	0.144 (0.067)	0.958 (0.025)	38	1	0	0	0.066 (0.016)	0.946 (0.014)
10	0	1	0	0.119 (0.034)	0.895 (0.020)	39	1	0	0	0.066 (0.016)	0.949 (0.015)
11	0	1	0	0.119 (0.034)	0.854 (0.016)	40	1	0	0	0.066 (0.016)	0.909 (0.014)
12	0	1	1	0.144 (0.067)	0.879 (0.025)	41	1	0	0	0.066 (0.016)	0.909 (0.014)
13	0	1	0	0.119 (0.034)	0.836 (0.014)	42	1	0	0	0.066 (0.016)	0.969 (0.018)
14	1	1	0	0.088 (0.031)	0.840 (0.033)	43	1	0	0	0.066 (0.016)	0.946 (0.014)
15	0	1	1	0.144 (0.067)	0.841 (0.014)	44	1	0	0	0.066 (0.016)	0.963 (0.014)
16	0	1	0	0.119 (0.034)	0.959 (0.017)	45	1	0	0	0.066 (0.016)	0.992 (0.007)
17	1	1	0	0.088 (0.031)	0.976 (0.015)	46	1	0	0	0.066 (0.016)	0.949 (0.014)
18	0	1	0	0.119 (0.034)	0.977 (0.020)	47	1	0	0	0.066 (0.016)	0.944 (0.014)
19	0	1	0	0.119 (0.034)	0.838 (0.015)	48	1	0	0	0.066 (0.016)	0.911 (0.013)
20	0	1	0	0.119 (0.034)	0.854 (0.015)	49	1	0	0	0.066 (0.016)	0.954 (0.017)
21	0	1	0	0.119 (0.034)	0.965 (0.016)	50	1	0	0	0.066 (0.016)	0.962 (0.013)
22	1	1	0	0.088 (0.031)	0.904 (0.015)	51	1	0	0	0.066 (0.016)	0.918 (0.017)
23	0	1	0	0.119 (0.034)	0.918 (0.027)	52	1	0	0	0.066 (0.016)	0.950 (0.013)
24	1	1	1	0.105 (0.053)	0.885 (0.015)	53	1	0	0	0.066 (0.016)	0.870 (0.016)
25	0	0	0	0.093 (0.025)	0.894 (0.013)	54	1	0	0	0.066 (0.016)	0.934 (0.014)
26	0	0	0	0.093 (0.025)	0.982 (0.013)	55	1	0	0	0.066 (0.016)	0.985 (0.012)
27	0	0	0	0.093 (0.025)	0.834 (0.014)	56	1	0	0	0.066 (0.016)	0.996 (0.004)
28	0	0	1	0.116 (0.063)	0.930 (0.025)	57	1	0	0	0.066 (0.016)	0.995 (0.005)
29	0	0	0	0.093 (0.025)	0.936 (0.016)	58	1	0	0	0.066 (0.016)	0.768 (0.016)
Average for branches with $s_{i2} = 1$, $s_{i3} = 1$ and $s_{i4} = 1$											0.890 (0.026)
Average for branches with $s_{i2} = 0$, $s_{i3} = 0$ and $s_{i4} = 0$											0.930 (0.014)

i	s_{i2}	s_{i3}	s_{i4}	λ_i	r_i	i	s_{i2}	s_{i3}	s_{i4}	λ_i	r_i
Average for branches with s_{i2} = 1, s_{i3} = 0 and s_{i4} = 0											0.934 (0.013)
Average for branches with s_{i2} = 0, s_{i3} = 1 and s_{i4} = 0											0.901 (0.021)
Average for branches with s_{i2} = 0, s_{i3} = 0 and s_{i4} = 1											0.930 (0.025)
Average for branches with s_{i2} = 1, s_{i3} = 1 and s_{i4} = 0											0.906 (0.021)
Average for branches with s_{i2} = 0, s_{i3} = 1 and s_{i4} = 1											0.894 (0.021)
Average for branches with s_{i2} = 1, s_{i3} = 0 and s_{i4} = 1											–
Average for branches with s_{i2} = 1											0.928 (0.015)
Average for branches with s_{i2} = 0											0.909 (0.019)
Average for branches with s_{i3} = 1											0.899 (0.021)
Average for branches with s_{i3} = 0											0.933 (0.014)
Average for branches with s_{i4} = 1											0.898 (0.023)
Average for branches with s_{i4} = 0											0.922 (0.016)
Average for all branches											0.919 (0.017)

Source: Authors' calculations.

It is important to note that our inference on individual efficiency levels is not sensitive to prior assumptions. Taking $r^* = 0.9$ (instead of 0.7) leads to only slightly higher posterior means (0.922 on average, instead of 0.919) and almost the same ranking of branches. The correlation coefficient between the individual posterior means (for $r^* = 0.7$ and $r^* = 0.9$) is 0.99978, and the Spearman rank correlation coefficient is 0.99951.

While inference on individual efficiency is insensitive to changes in r^*, values of this prior hyperparameter exert influence on the posterior results for the parameters $\gamma_j = \ln(\phi_j)$, which parameterize the sampling mean λ_i of inefficiency (individual effect) z_i; see (2). Table 5 shows the posterior means and standard deviations of γ_j's under three very different values of r^* (one of them, 0.5, is too low to be reasonable). The positive, although decreasing with r^*, posterior mean of γ_2 would suggest that "depository" branches ($s_{i2} = 1$) tend to be more efficient. The negative (decreasing with r^*) posterior means of γ_3 and γ_4 would mean that large branches ($s_{i3} = 1$) and branches that have subbranches ($s_{i4} = 1$) tend to be less efficient. The posterior means of γ_j (j = 2, 3, 4) confirm our initial conjectures. This, however, should be interpreted with caution as the posterior standard deviations of γ_j are very large. In order to test possible systematic differences in cost efficiency, we use the Bayesian Lindley type test based on Highest Posterior Density (HPD) regions.

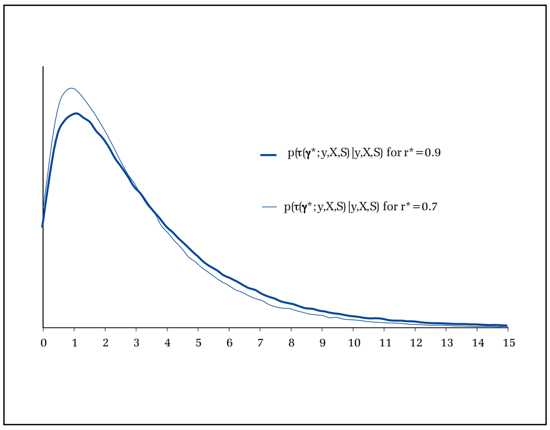
Since the marginal posterior distribution of $\boldsymbol{\gamma}^* = (\gamma_2\ \gamma_3\ \gamma_4)'$ is approximately Normal with mean $E(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{X}, \mathbf{S})$ and covariance matrix $V(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{X}, \mathbf{S})$, the quadratic form $\tau(\boldsymbol{\gamma}^*; \mathbf{y}, \mathbf{X}, \mathbf{S}) = [\boldsymbol{\gamma}^* - E(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{X}, \mathbf{S})]' V^{-1} (\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{X}, \mathbf{S}) [\boldsymbol{\gamma}^* - E(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{X}, \mathbf{S})]$ has the posterior distribution close to the chi-square distribution with 3 degrees of freedom. Figure 8 presents the exact posterior density of $\tau(\boldsymbol{\gamma}^*; \mathbf{y}, \mathbf{X}, \mathbf{S})$, obtained as a by-product of the Gibbs sampler for both $r^* = 0.7$ and $r^* = 0.9$. The tested value of $\boldsymbol{\gamma}^*$, i.e. $\mathbf{0}$, leads to $\tau(\mathbf{0}; \mathbf{y}, \mathbf{X}, \mathbf{S})$ equal to 3.23 under $r^* = 0.7$ and to 3.34 under $r^* = 0.9$. In both cases there is no reason to reject $\boldsymbol{\gamma}^* = \mathbf{0}$ as $\tau(\mathbf{0}; \mathbf{y}, \mathbf{X}, \mathbf{S})$ lies in HPD intervals of probability content at least 0.65. This exact Bayesian counterpart of the approximate chi-square test shows that none of the variables s_{ij} (j = 2, 3, 4) introduced in our VED specification helps in explaining differences in individual short-run cost

Table 5. *Posterior means and standard deviations of* γ *(VED with* m = 4*)*

	$r^* = 0.5$		$r^* = 0.7$		$r^* = 0.9$	
	E(·|y,X,S)	D(·|y,X,S)	E(·|y,X,S)	D(·|y,X,S)	E(·|y,X,S)	D·|y,X,S)
γ_1	2.186	0.250	2.411	0.261	2.639	0.278
γ_2	0.475	0.269	0.327	0.276	0.171	0.285
γ_3	-0.105	0.308	-0.247	0.313	-0.397	0.319
γ_4	-0.133	0.418	-0.135	0.415	-0.146	0.413

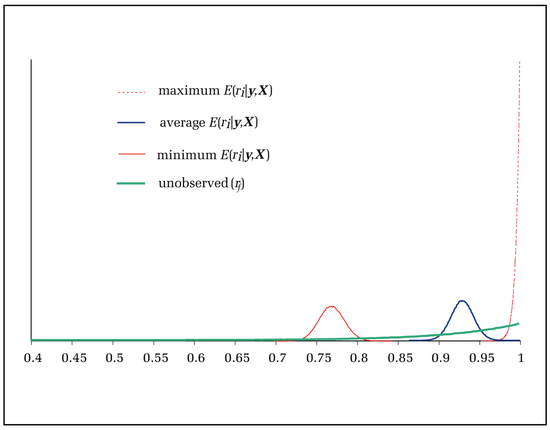
Source: Authors' calculations.

Figure 8. *Posterior densities of $\tau(\gamma^*; \mathbf{y}, \mathbf{X}, \mathbf{S})$ for $r^* = 0.7$ and $r^* = 0.9$*



Source: Authors' calculations.

Figure 9. *Marginal posterior densities of cost efficiency r_i (CED, $r^* = 0.7$)*



Source: Authors' calculations.

efficiency. Thus, these differences can be considered random and the simpler CED specification ($m = 1$, $s_{i1} = 1$) can be adopted. In fact, the CED model leads to very similar posterior results on technology and individual cost efficiency.

The CED specification treats all individual effects z_i (given the parameters of the model) as independent drawings from the same Exponential distribution with mean $\lambda = 1/\phi_1$; see Osiewalski (2001), chapter 7. Using the posterior density of λ (with mean 0.086 and standard deviation 0.015) we can integrate this parameter out and obtain the marginal posterior distribution of efficiency of an unobserved branch (the predictive distribution for individual efficiency):

$$p(r_f|\mathbf{y},\mathbf{X}) = r_f^{-1}\int_0^\infty f_G\left(-\ln(r_f)\middle|1,\lambda^{-1}\right)p(\lambda|\mathbf{y},\mathbf{X})d\lambda \qquad (17)$$

approximated (using the Gibbs sampler) by

$$p(r_f|\mathbf{y},\mathbf{X}) \approx r_f^{-1}\frac{1}{M}\sum_{l=1}^{M} f_G\left(-\ln(r_f)\middle|1,\left(\lambda^{(l)}\right)^{-1}\right)$$

This rather diffuse distribution, presented in Figure 9, gives the overall picture of the short-run cost efficiency of the analysed branches. Its mean, 0.921, is the same as the simple average of individual posterior means (0.921) but its standard deviation is very large (0.075). Thus, the posterior distribution of r_f covers results on efficiency for all branches – from the least to the most efficient. Figure 9 also presents the marginal posterior densities $p(r_i|\ \mathbf{y}, \mathbf{X})$ for the branches with the maximum, minimum and average posterior means of r_i. These densities are quite sharp as we use panel data and efficiencies are treated as individual effects.

6. Conclusion

In this paper we have reviewed the Bayesian analysis of stochastic frontier models, arguing that Gibbs sampling can be used to greatly reduce the computational burden inherent to this analysis. Following KOS (1994b; 1997), we have shown how the posterior conditional densities can be used to set up a Gibbs sampler in the case of inefficiencies treated as individual effects. The structure of the Gibbs sampler follows naturally from viewing the inefficiency terms as additional parameters in a regression model; see Fernández et al. (1997). In important special cases all conditionals are Gamma or truncated Normal distributions, which leads to enormous computational gains.

We have applied the Bayesian methodology to make posterior inference on the technology and short-run cost efficiency of 58 branches of a Polish bank. Our results, based on panel data from 4 quarters of one year and a translog variable cost frontier, indicate increasing returns to scale (varying with the branch output level) and no systematic differences in efficiency that could be explained by the three dummy variables under consideration. The example also illustrates that cooperation with Bayesian econometricians may create important insights into the economic functioning of the bank. The management may learn not only about the basic microeconomic characteristics of each branch, but also about the branch efficiency and its possible determinants.

Our cost model has been formulated in terms of one aggregate product, Q, but extensions to more products are straightforward. Marzec (2000) and Marzec, Osiewalski (2001) present posterior inference for the case where Q is split into two categories: commercial loans and other products. The basic results on technology and efficiency remain unchanged (with respect to the case of one aggregate product) but, of course, inference on scope economies is also possible at almost negligible additional computational cost. Since, as we argue in our other work, inference on scope economies or effects of specialisation requires new measures, we have not discussed these issues in the present paper, which is focused mainly on the use of Bayesian statistical methodology in cost efficiency analysis for the banking sector.

References

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

Akhaiven J., Swamy P.A.V.B., Taubman S.B., Singamsetti R.N (1997), *A general method of deriving the inefficiencies of bank from a profit function*, "Journal of Productivity Analysis", Vol. 8, pp. 71–93.

Altunbas Y., Liu M., Molyneux P., Seth R. (2000), *Efficiency and risk in Japanese banking*, "Journal of Banking and Finance", Vol. 24, pp. 1605–1628.

Bauer W.B., Hancock D. (1993), *The efficiency of the Federal Reserve in providing check processing services*, "Journal of Banking and Finance", Vol. 17, pp. 287–311.

Berger A.N., Leusner J.H., Mingo J.J. (1997), *The efficiency of bank branches*, "Journal of Monetary Economics", Vol. 40, pp. 141–162.

Berger A.N., Mester L. (1997), *Inside the black box: What explains differences in the efficiencies of financial institutions?*, "Journal of Banking and Finance", Vol. 21, pp. 895-947.

Berger A.N., DeYoung R. (1997), *Problem loans and cost efficiency in commercial banks*, "Journal of Banking and Finance", Vol. 21, pp. 849–870.

van den Broeck J., Koop G., Osiewalski J., Steel M.F.J. (1994), *Stochastic frontier models: A Bayesian perspective*, "Journal of Econometrics", Vol. 61, pp. 273–303.

Casella G., George E. (1992), *Explaining the Gibbs sampler*, "The American Statistician", Vol. 46, pp. 167–174.

Cebenoyan A.S., Cooperman E.S., Register C.A., Hudgins S.A. (1993), *The relative efficiency of Stock versus Mutual S&Ls: A stochastic cost frontier approach*, "Journal of Financial Services Research", Vol. 7, pp. 151–170.

Christensen L.R., Greene W.H. (1976), *Economies of scale in U.S. electric power generation*, "Journal of Political Economy", Vol. 84, pp. 655–676.

English M., Grosskopf S., Hayes K., Yaiswarng S. (1993), *Output allocative and technical efficiency of banks*, "Journal of Banking and Finance", Vol. 17, pp. 349–366.

Fernández C., Osiewalski J., Steel M.F.J. (1997), *On the use of panel data in stochastic frontier models with improper priors*, "Journal of Econometrics", Vol. 79, pp. 169–193.

Ferrier G.D., Lovell C.A.K. (1990), *Measuring cost efficiency in banking: econometric and linear programming evidence*, "Journal of Econometrics", Vol. 46, pp. 229–245.

Grabowski R., Ragan N., Rezvanian R. (1993), *Organizational forms in banking: An empirical investigation of cost efficiency*, "Journal of Banking and Finance", Vol. 17, pp. 531–538.

Hassan Y.A., Grabowski R., Pasurka C., Ragan N. (1990), *Technical, scale, and allocative efficiencies in U.S. banking: An empirical investigation*, "Review of Economics and Statistics", Vol. 72, pp. 211–218.

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

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

Koop G., Osiewalski J., Steel. M.F.J. (1994a), *Bayesian efficiency analysis with a flexible form: The AIM cost function*, "Journal of Business and Economic Statistics", Vol. 12, pp. 339–346.

Koop G., Osiewalski J., Steel. M.F.J. (1994b), *Hospital efficiency analysis through individual effects: A Bayesian approach*, "CentER Discussion Paper", No. 9447, Tilburg University, Tilburg.

Koop G., Osiewalski J., Steel M.F.J. (1997), *Bayesian efficiency analysis through individual effects: Hospital cost frontiers*, "Journal of Econometrics", Vol. 76, pp. 77–105.

Koop G, Osiewalski J., Steel M.F.J. (1999), *The components of output growth: A stochastic frontier analysis*, "Oxford Bulletin of Economics and Statistics", Vol. 61, pp. 455–487.

Koop G, Osiewalski J., Steel M.F.J. (2000a), *Modeling the sources of output growth in a panel of countries*, "Journal of Business and Economic Statistics", Vol. 18, pp. 284–299.

Koop G, Osiewalski J., Steel M.F.J. (2000b), *A stochastic frontier analysis of output level and growth in Poland and Western economies*, "Economics of Planning", Vol. 33, pp. 185–202.

Koop G., Steel M.F.J., Osiewalski J. (1995), *Posterior analysis of stochastic frontier models using Gibbs Sampling*, "Computational Statistics", Vol. 10, pp. 353–373.

Kraft E., Tirtiroglu D. (1998), *Bank efficiency in Croatia: A stochastic-frontier analysis*, "Journal of Comparative Economics", Vol. 26, pp. 282–300.

Kumbhakar S.C. (1997), *Modeling allocative inefficiency in a translog cost function and cost share equations: An exact relationship*, "Journal of Econometrics", Vol. 76, pp. 351–356.

Marzec J. (2000), *Ekonometryczna analiza efektywności kosztów w bankach komercyjnych* [Econometric analysis on cost efficiency in commercial banks], unpublished doctoral dissertation (in Polish), Cracow University of Economics, Kraków.

Marzec J., Osiewalski J. (2001), *Funkcja kosztu dla oddziałów banku: mierniki korzyści specjalizacji*, "Prace Naukowe Akademii Ekonomicznej we Wrocławiu", No. 895, pp. 155–165.

Marzec J., Osiewalski J. (2003), *Bayesowskie graniczne modele kosztów dla oddziałów banku. Wnioskowanie o efektywności kosztowej i jej determinantach*, "Zeszyty Naukowe Akademii Ekonomicznej w Krakowie", nr 628, pp. 23–51.

Meeusen W., van den Broeck J. (1997), *Efficiency estimation from Cobb-Douglas production functions with composed error*, "International Economic Review", Vol. 8, pp. 435–444.

Mielnik M., Ławrynowicz M. (2002), *Badanie efektywności technicznej banków komercyjnych w Polsce metodą DEA*, "Bank i Kredyt", No. 5, pp. 52–64.

Mester L.J. (1987), *A multiproduct cost study of savings and loans*, "Journal of Finance", Vol. 42, pp. 423–445.

Mester L.J. (1993), *Efficiency in the savings and loan industry*, "Journal of Banking and Finance", Vol. 17, pp. 267–286.

Mester L.J. (1997), *Measuring efficiency at U.S. banks: Accounting for heterogeneity is important*, "European Journal of Operational Research", Vol. 98, pp. 230-242.

Muldur U., Sassenou M. (1993), *Economies of scale and scope in French banking and savings institutions*, "Journal of Productivity Analysis", Vol. 4, pp. 51–72.

Nerlove M. (1963), *Return to scale in electricity supply*, in: C.F Christ (ed.), *Measurement in Economics*, Stanford University Press, Stanford.

Noulas A.G., Subhash C.R., Miller S.M. (1990), *Returns to scale and input substitution for large U.S. banks*, "Journal of Money, Credit, and Banking", Vol. 22, pp. 94–108.

O'Hagan A. (1994), *Bayesian Inference*, Edward Arnold, London.

Osiewalski J. (2001), *Ekonometria bayesowska w zastosowaniach* [Bayesian econometrics in applications], Publishing house of the Cracow University of Economics, Cracow.

Osiewalski J., Marzec J. (1998), *Bayesian analysis of cost efficiency with an application to bank branches*, in: E. Miklaszewska (ed.), *Global Tendencies and Changes in East European Banking*, Jagiellonian University, Cracow.

Osiewalski J., Steel M.F.J. (1998), *Numerical tools for the Bayesian analysis of stochastic frontier models*, "Journal of Productivity Analysis", Vol. 10, pp. 103–117.

Pawłowska M. (2003a), *Wpływ fuzji i przejęć na efektywność w sektorze banków komercyjnych w Polsce w latach 1997–2001*, "Bank i Kredyt", nr 2, pp. 20–34.

Pawłowska M. (2003b), *Wpływ zmian w strukturze polskiego sektora bankowego na jego efektywność w latach 1997-2002 (podejście nieparametryczne)*, "Bank i Kredyt", No. 11–12, pp. 51–64.

Pitt M.M., Lee L.F. (1981), *The measurement and sources of technical inefficiency in the Indonesian weaving industry*, "Journal of Development Economics", Vol. 9, pp. 43–64.

Ritter C. (1993), *The Normal-Gamma frontier model under a common vague prior does not produce a proper posterior*, Manuscript, Université Catholique de Louvain, Louvain-la-Neuve.

Rogers K.E. (1998), *Nontraditional activities and the efficiency of US commercial banks*, "Journal of Banking and Finance", Vol. 22, pp. 467–482.

Salvanes K.G., Tjøtta S. (1998), *A note on the importance of testing for regularities for estimated flexible functional forms*, "Journal of Productivity Analysis", Vol. 9, pp. 133–143.

Schmidt P., Sickles R.C. (1984), *Production frontiers and panel data*, "Journal of Business and Economic Statistics", Vol. 2, pp. 367–374.

Sealey C.W., Lindley J.T. (1977), *Inputs, outputs, and a theory of production and cost at depository financial institutions*, "Journal of Finance", Vol. 32, pp. 1251–1266.

Simon C.P., Blume L. (1994), *Mathematics for Economists*, W.W. Norton, New York.

Tierney L. (1994), *Markov chains for exploring posterior distributions (with discussion)*, "Annals of Statistics", Vol. 22, pp. 1701–1762.

Zardokoohi A., Kolari J. (1994), *Branch office economies of scale and scope: Evidence from savings banks in Finland*, "Journal of Banking and Finance", Vol. 18, pp. 421–432.