Empirical Economics manuscript No.
(will be inserted by the editor)

Addressing Endogeneity when Estimating Stochastic Ray Production Frontiers: A Bayesian Approach

Mike Tsionas · Marwan Izzeldin ·
Arne Henningsen · Evaggelos Paravalos

Received: date / Accepted: date

Abstract We propose a Bayesian approach for inference in the stochastic ray production frontier (SRPF), which can model multiple-input–multiple-output production technologies even in case of zero output quantities, i.e., if some outputs are not produced by some of the firms. However, the econometric estimation of the SRPF—as the estimation of distance functions in general—is susceptible to endogeneity problems. To address these problems, we apply a profit-maximizing framework to derive a system of equations after incorporating technical inefficiency. As the latter enters non-trivially into the system of equations and as the Jacobian is highly complicated, we use Monte Carlo methods of inference. Using US banking data to illustrate our innovative approach, we also address the problems of missing prices and the dependence on the ordering of the outputs via model averaging.

Keywords Stochastic ray production frontier · Technical inefficiency · Endogeneity · Bayesian inference · Model averaging

JEL codes C11 · C13 · D24

The authors wish to thank Subal C. Kumbhakar and two anonymous reviewers for valuable discussions and comments. E. Paravalos gratefully acknowledges financial support from the General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI).

Mike Tsionas
Department of Economics, Lancaster University Management School, LA1 4YX, UK,
E-mail: m.tsionas@lancaster.ac.uk

Marwan Izzeldin
Department of Economics, Lancaster University Management School, LA1 4YX, UK, E-mail: m.izzeldin@lancaster.ac.uk

Arne Henningsen
Department of Food and Resource Economics, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg C, Denmark, Tel.: +45-514396180, E-mail: arne@ifro.ku.dk

Evaggelos Paravalos
Department of Economics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece, E-mail: paravalosev@aueb.gr

1 Introduction

The stochastic frontier production model, originally introduced by Meeusen and van den Broeck (1977), Aigner et al. (1977), and Battese and Corra (1977), incorporates a composed error term consisting of a symmetric error with zero mean representing statistical noise and a non-negative error expressing technical inefficiency. A significant limitation of the original stochastic production frontier model is that it cannot adequately handle the case of multiple outputs. In the case of multiple-input, multiple-output technologies, the standard approach has been to take advantage of the dual approach and estimate instead either the cost frontier function (e.g., Ferrier and Lovell 1990; Koop et al. 1997; Rosko 2001; Filippini and Farsi 2004; Orea and Kumbhakar 2004; Huang and Wang 2004) or the profit frontier function (e.g., Kumbhakar and Bhattacharyya 1992; Vivas 1997; Berger and Mester 1997; Humphrey and Pulley 1997; Akhavein et al. 1997; Kumbhakar and Lovell 2000; Kumbhakar 2001).[1] However, this requires the availability of input and output prices as well as data on cost or profits and the behavioral postulate of either cost minimization or profit maximization, which may not be appropriate assumptions for many empirical analyses. If any of the above requirements fails, duality cannot be exploited and a primal approach would need to be adopted.

The framework of a primal approach based on stochastic distance functions was developed in the literature to address those criticisms. Originally, Färe et al. (1993) used the concept of an output distance function, originally introduced by Shephard (1970), to model multiple-output technologies. Lovell et al. (1994) developed a multiple-output generalization of the primal approach and proposed a transformation of the data, while making use of the linear homogeneity property of the distance function to overcome the problem of the unobserved distance measure. The main contribution of distance functions is that they allow the specification of multiple-input, multiple-output technologies when information about prices may not be available or when the cost or profit representations do not present a preferable alternative due to violation of the related behavioral assumptions (Coelli and Perelman 1999, 2000; Sickles et al. 2002; Atkinson and Dorfman 2005). A considerable disadvantage of the distance function approach is that the regressors are possibly endogenous and, thus, the estimates are possibly inconsistent due to aspects of inherent simultaneous equations as discussed in Grosskopf et al. (1997) and Cuesta and Orea (2002).

Löthgren (1997) developed a multiple-output generalization of the single-output stochastic frontier production model, called the stochastic ray production frontier (SRPF) model, which defines the Euclidean norm of the vector of output quantities as a function of input quantities and polar-coordinate angles of the output quantities. The SRPF—which describes the maximum level of the Euclidean norm of the output vector that can be achieved for a given level

[1] For a presentation of multiple-output production and duality theory, see Färe and Primont (1995, 1996).

of inputs and a given output mix under the technology—offers an alternative primal approach of estimation for stochastic multiple-output frontier functions (e.g., Gerdtham et al. 1999; Löthgren 2000; Niquidet and Nelson 2010; Bhattacharyya and Pal 2013). The issues of endogeneity and inconsistency, prevalent in stochastic distance functions, are less profound in the SRPF as the error terms affect outputs radially given the exogenous output mix.

As it is commonly the case with distance functions, the explanatory variables of the SRPF—consisting of input quantities and angles between the output quantities—can be correlated with the inefficiency term and/or the noise term. Thus, estimation of the SRPF via classical econometric techniques (e.g., Aigner et al. 1977; Pitt and Lee 1981; Jondrow et al. 1982; Schmidt and Sickles 1984; Kumbhakar 1990; Battese and Coelli 1992; Kumbhakar and Heshmati 1995; Greene 2005) may give inconsistent estimates (Kumbhakar and Lovell 2000). In addressing the susceptibility to endogeneity problems, we contribute to the Bayesian efficiency analysis of stochastic frontier models (e.g., van den Broeck et al. 1994; Koop et al. 1994, 1995; Tsionas 2000, 2002, 2006; Koop and Steel 2001) by combining the SRPF with the first-order conditions of the related profit-maximizing framework to address potential endogeneity problems. In addition, we propose a Bayesian procedure of inference for the parameters of the SRPF and technical inefficiency using model averaging to address the dependence of the SRPF upon the ordering of the outputs.[2]

2 Microeconomic specification and first-order conditions

The SRPF, originally proposed by Löthgren (1997) and recently improved by Henningsen, Bělín, and Henningsen (2017), is a non-standard representation of an output distance function (Henningsen et al. 2015). Suppose $x \in \Re^K$ denotes a vector of inputs and $y \in \Re^M$ represents a vector of outputs, Henningsen et al. (2017) have shown that a Shephard (radial) output distance function (ODF), i.e., $D(x, y) = \min\left\{\lambda > 0 : (y\lambda^{-1}, x) \text{ can be produced}\right\}$ with $0 < D(x, y) \leq 1$,

[2] Our approach invokes the same assumption as dual approaches, i.e., availability of input (and output) prices and the assumption of profit maximization (or at least cost minimization). However, in certain empirical applications, our approach has the following advantages: (a) Dual approaches require notable variation in input (and output) prices between observations. However, if transaction costs are low and markets are working well, the "law of one price" approximately holds, which leaves too little variation between observations for estimating a dual approach. By our alternative approach, there is no requirement for variation in input (and output) prices. (b) To obtain estimates of output-oriented technical efficiency, a dual approach is (at best) very complicated to implement. (c) Although, in theory, primal and dual approaches should give the same result, in practice this is not assured. Hence, the preference is for a primal approach, as presented in this paper.

can be transformed to a SPRF specified as:[3]

$$\ln D(x, y) = \ln ||y|| - F(\ln x, \vartheta(y)) \tag{1}$$

$$\ln ||y|| = F(\ln x, \vartheta(y)) - u, \tag{2}$$

using the polar-coordinate representation of the output vector $y = ||y|| \cdot m(\vartheta)$, where $||y|| = \sqrt{\sum_{m=1}^{M} y_m^2}$, $\vartheta(y)$ with $\vartheta_m(y) = \arccos\left(y_m \Big/ \sqrt{\sum_{j=m}^{M} y_j^2}\right)$ $\forall m = 1, ..., M-1$, and $m(\vartheta) : [0, \pi/2]^{M-1} \rightarrow [0, 1]^M$ represent the Euclidean norm (length), the polar-coordinate angles, and the transformation function of the angles to the mix, respectively, of the output vector y, and $u = -\ln D(x, y) \geq 0$ represents output-oriented technical inefficiency. Adding a time trend t to take into account technical change, assuming a quadratic functional form of $F(\cdot)$, and finally adding a two-sided error term v to take into account statistical noise, we get a Translog SRPF:[4]

$$\begin{aligned} \ln ||y|| = {} & \alpha_0 + \sum_{m=1}^{M-1} \alpha_m \vartheta_m + \tfrac{1}{2} \sum_{m=1}^{M-1} \sum_{m'=1}^{M-1} \alpha_{mm'} \vartheta_m \vartheta_{m'} \\ & + \sum_{k=1}^{K} \beta_k \ln x_k + \tfrac{1}{2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \beta_{kk'} \ln x_k \ln x_{k'} \\ & + \sum_{m=1}^{M-1} \sum_{k=1}^{K} \gamma_{mk} \ln x_k \vartheta_m + \zeta_{\mathfrak{T}} t + \tfrac{1}{2} \zeta_{\mathfrak{T}\mathfrak{T}} t^2 \\ & + \sum_{m=1}^{M-1} \zeta_{\mathfrak{T}m}^{y} t\, \vartheta_m + \sum_{k=1}^{K} \zeta_{\mathfrak{T}k}^{x} t\, \ln x_k + v - u, \end{aligned} \tag{3}$$

which can be estimated as a stochastic frontier model.

Similarly to Tsionas et al. (2015) who take into account endogeneity by estimating an input distance function along with cost-minimizing first-order conditions (FOC), we estimate an SRPF (3) along with the corresponding profit-maximizing FOC. We derive the FOC from the profit maximization problem:

$$\max_{x \in \Re_K^+, y \in \Re_M^+} p'y - w'x, \text{ s.t. } D(x, y) = e^{-u}, \tag{4}$$

[3] We refer the reader to Löthgren (1997), Gerdtham et al. (1999) and Löthgren (2000) for more details on the SRPF. The traditional specification of a Translog Shephard output distance function that corresponds to the stochastic-ray output distance function specified in equation (1) is $\ln D(x, y) = \ln y_m + F(\ln x, \ln(y/y_m))$, where $F(\cdot)$ denotes a quadratic functional form and m with $1 \leq m \leq M$ indicates an arbitrarily chosen output that is used as numéraire to impose linear homogeneity in output quantities. Hence, the traditional Translog specification fulfils linear homogeneity through an arbitrarily chosen output and implicitly assumes that $\ln(D(x, y))$ is a quadratic function of the logarithms of the input quantities and the logarithms of the normalised output quantities, while the Translog SRPF functional form fulfils linear homogeneity through the length of the vector of output quantities and implicitly assumes that $\ln(D(x, y))$ is a quadratic function of the logarithms of the input quantities and the angles of the vector of output quantities.

[4] An advantage of not taking logarithms of the angles (unlike the specification in Löthgren 1997) is that this specification can handle zero output quantities (Henningsen et al. 2017).

where technical inefficiency u is taken as given to the producer. If λ denotes the Lagrange multiplier, the FOC are:

$$w_k = -\lambda \frac{\partial D(x,y)}{\partial x_k} = -\lambda \frac{\partial \ln D(x,y)}{\partial \ln x_k} \frac{D(x,y)}{x_k} \; \forall \, k = 1, ..., K, \tag{5}$$

$$p_m = \lambda \frac{\partial D(x,y)}{\partial y_m} = \lambda \frac{\partial \ln D(x,y)}{\partial \ln y_m} \frac{D(x,y)}{y_m} \; \forall \, m = 1, ..., M, \tag{6}$$

where $w_k; k = 1, ..., K$ denotes the price of the kth input and $p_m; m = 1, ..., M$ denotes the price of the mth output. As the Lagrange multiplier λ is equal to total revenue at full efficiency (Brümmer et al. 2002), i.e., $\lambda = p'y$ for $D(x,y) = 1$ and, thus, $\lambda = p' \left(y/D(x,y)\right)$ for $0 < D(x,y) \leq 1$, we can eliminate λ from the FOC and re-arrange them to get:

$$\frac{w_k x_k}{p'y} = -\frac{\partial \ln D(x,y)}{\partial \ln x_k} \; \forall \, k = 1, ..., K, \tag{7}$$

$$\frac{p_m y_m}{p'y} = \frac{\partial \ln D(x,y)}{\partial \ln y_m} \; \forall \, m = 1, ..., M. \tag{8}$$

The right-hand sides of equations (7) and (8) are the distance elasticities of the inputs and outputs, respectively, that can be calculated as:[5]

$$\frac{\partial \ln D(\ln x, \vartheta)}{\partial \ln x_k} = -\frac{\partial F(\ln x, \vartheta)}{\partial \ln x_k} \tag{9}$$

$$= -\beta_k - \sum_{k'=1}^{K} \beta_{kk'} \ln x_{k'} - \sum_{m=1}^{M-1} \gamma_{mk} \vartheta_m - \zeta^x_{\mathfrak{T}k} t \tag{10}$$

$$\forall \, k = 1, ..., K,$$

$$\frac{\partial \ln D(\ln x, \vartheta)}{\partial \ln y_m} = \frac{\partial \ln ||y||}{\partial \ln y_m} - \frac{\partial F(\ln x, \vartheta)}{\partial \ln y_m} \tag{11}$$

$$= \frac{y_m^2}{||y||^2} - \sum_{m'=1}^{\min(m,M-1)} \frac{\partial F(\ln x, \vartheta)}{\partial \vartheta_{m'}} \cdot \frac{\partial \vartheta_{m'}}{\partial \ln y_m} \tag{12}$$

$$= \frac{y_m^2}{||y||^2} + \sum_{m'=1}^{\min(m,M-1)} \left\{ \left(\alpha_{m'} + \sum_{m^*=1}^{M-1} \alpha_{m'm^*} \vartheta_{m^*} + \sum_{k=1}^{K} \gamma_{m'k} \ln x_k + \zeta^y_{\mathfrak{T}m'} t \right) \frac{y_m S_{m'} \left(\delta_{mm'} - y_m y_{m'} S^2_{m'}\right)}{\sqrt{1 - y^2_{m'} S^2_{m'}}} \right\} \; \forall \, m = 1, ..., M, \tag{13}$$

where $S_{m'} = 1 \Big/ \sqrt{\sum_{m^*=m'}^{M} y^2_{m^*}}$ and $\delta_{mm'} = 1$ if $m = m'$ and zero otherwise. The monotonicity conditions are: $\partial \ln D(x,y)/\partial \ln x_k \leq 0 \; \forall \, k = 1, ..., K$ and

[5] Note that $\partial \vartheta_{m'} / \partial \ln y_m = (\partial \vartheta_{m'} / \partial y_m) \, y_m$, $\partial \vartheta_{m'} / \partial \ln y_m = 0 \; \forall \, m < m'$, and $\partial \arccos(z)/\partial z = -1/\sqrt{1 - z^2}$.

$\partial \ln D(x,y)/\partial \ln y_m \geq 0 \;\forall\; m = 1,...,M$ (see Kumbhakar and Lovell 2000, p. 32).

Based on the estimated parameters, we can calculate the elasticity of scale $RTS_{it} = -\sum_{k=1}^{K} \partial \ln D_{it}/\partial \ln x_{k,it}$ (Färe and Grosskopf 1994, p. 103), the annual rate of technical change $TC_{it} = \zeta_{\mathfrak{T}} + \zeta_{\mathfrak{T}\mathfrak{T}} t + \sum_{m=1}^{M-1} \zeta^{y}_{\mathfrak{T}m} \vartheta_{m,it} + \sum_{k=1}^{K} \zeta^{x}_{\mathfrak{T}k} \ln x_{k,it}$, the annual efficiency change $EC_{it} = e^{-u_{it}} - e^{-u_{it-1}}$, the annual scale efficiency change $SEC_{it} = -\left(RTS_{it} - 1\right) \sum_{k=1}^{K} \left(\partial \ln D_{it}/\partial \ln x_{k,it}\right) \left(\ln x_{k,it} - \ln x_{k,it-1}\right)/RTS_{it}$, and the annual rate of productivity growth $PG_{it} = TC_{it} + EC_{it} + SEC_{it}$.

3 Econometric specification and estimation procedure

In the above formulation, we have $K + M$ endogenous variables $(x_1, \ldots, x_K, y_1, \ldots, y_M$ or $x_1, \ldots, x_K, ||y||, \vartheta_1, \ldots, \vartheta_{M-1})$ but $K + M + 1$ equations (2, 7, 8). However, given that the revenue shares $(p_m y_m / p'y)$ and—due to the linear homogeneity of output distance functions—the distance elasticities of the outputs $(\partial \ln D(x,y)/\partial \ln y_m)$ both sum up to one, i.e., $\sum_{m=1}^{M} \left(p_m y_m\right)/\left(p'y\right) = \sum_{m=1}^{M} \partial \ln D(x,y)/\partial \ln y_m = 1$, one of the FOC regarding the outputs (8) is redundant and, thus, needs to be dropped from the estimation to avoid singularity. Given that the removal of one of these equations does not remove any information, the estimation results are invariant to the equation that is removed (Barten 1969). Hence, the system of equations that we estimate contains $K+M$ equations in total: the SRPF (2), K FOC with respect to the input quantities (7), and $(M-1)$ FOC with respect to the output quantities (8).

The system of equations used in the estimation, obtained after adding error terms to the FOC, for the case of panel data $(i = 1, ..., n,\; t = 1, ..., T)$ is:

$$
\begin{aligned}
v_{it} = {} & \ln ||y_{it}|| - \alpha_0 - \sum_{m=1}^{M-1} \alpha_m \vartheta_{m,it} - \tfrac{1}{2} \sum_{m=1}^{M-1} \sum_{m'=1}^{M-1} \alpha_{mm'} \vartheta_{m,it} \vartheta_{m',it} \\
& - \sum_{k=1}^{K} \beta_k \ln x_{k,it} - \tfrac{1}{2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \beta_{kk'} \ln x_{k,it} \ln x_{k',it} \\
& - \sum_{m=1}^{M-1} \sum_{k=1}^{K} \gamma_{mk} \ln x_{k,it} \vartheta_{m,it} - \zeta_{\mathfrak{T}} t + \tfrac{1}{2} \zeta_{\mathfrak{T}\mathfrak{T}} \mathrm{t}^2 \\
& - \sum_{m=1}^{M-1} \zeta^{y}_{\mathfrak{T}m} t\, \vartheta_{m,it} - \sum_{k=1}^{K} \zeta^{x}_{\mathfrak{T}k} t\; \ln x_{k,it} + u_{it},
\end{aligned} \tag{14}
$$

$$
\begin{aligned}
v^{x}_{k,it} = {} & \frac{w_{k,it} x_{k,it}}{p'_{it} y_{it}} - \beta_k - \sum_{k'=1}^{K} \beta_{kk'} \ln x_{k'} - \sum_{m=1}^{M-1} \gamma_{mk} \vartheta_m - \zeta^{x}_{\mathfrak{T}k} t \\
& \forall\; k = 1, ..., K,
\end{aligned} \tag{15}
$$

$$
v^{y}_{m,it} = \frac{p_{m,it} y_{m,it}}{p'_{it} y_{it}} - \frac{y^2_{m,it}}{||y_{it}||^2} \tag{16}
$$

$$-\sum_{m'=1}^{\min(m,M-1)}\left(\alpha_{m'}+\sum_{m^*=1}^{M-1}\alpha_{m'm^*}\vartheta_{m^*,it}+\sum_{k=1}^{K}\gamma_{m'k}\ln x_{k,it}+\zeta^{y}_{\mathfrak{T}m'}t\right)$$
$$\left(\frac{y_{m,it}S_{m',it}\left(\delta_{mm'}-y_{m,it}y_{m',it}S^2_{m',it}\right)}{\sqrt{1-y^2_{m',it}S^2_{m',it}}}\right)\ \forall\ m=1,...,M-1.$$

We denote the vector of error terms by $\mathbf{v}_{it}\equiv\left[\upsilon_{it},\upsilon^x_{1,it},...,\upsilon^x_{K,it},\upsilon^y_{1,it},\ldots,\upsilon^y_{M-1,it}\right]'$ and assume that it follows a $(K+M)$-variate normal distribution, i.e., $\mathbf{v}_{it}\sim\mathcal{N}_{K+M}\left(\mathbf{0},\boldsymbol{\Sigma}\right)\ \forall\ i=1,...,n;\ t=1,...,T$, where $\boldsymbol{\Sigma}=\operatorname{diag}\left(\sigma^2_\upsilon,\sigma^2_{\upsilon^x_1},...,\sigma^2_{\upsilon^x_K},\sigma^2_{\upsilon^y_1},...,\sigma^2_{\upsilon^y_{M-1}}\right)$. For technical inefficiency, we make the standard assumption that $u_{it}\sim\mathcal{N}^+\left(0,\sigma^2_u\right)$ independently of all error terms in $\mathbf{v}_{it}$ and all regressors. If we denote the vector of unknown parameters $(\alpha,\beta,\gamma,\zeta)$ by $\theta\in\Theta\subset\Re^D$, the system can be written compactly as follows:

$$\mathbf{v}_{it}=\mathbf{F}(\theta,u_{it};\mathcal{Y}_{it}),\tag{17}$$

where $\mathbf{F}$ is a $(K+M)$-dimensional function and $\mathcal{Y}_{it}\equiv(y'_{it},\ln x'_{it},w'_{it},p'_{it},t)'$ denotes the values of all variables at observation (i,t).[6]

Let $\mathcal{Y}=\{\mathcal{Y}_{it}\}$ denote the entire data set, the likelihood function of the system can be written in the form:

$$\begin{aligned}\mathcal{L}(\theta,\sigma_u,\boldsymbol{\Sigma};\mathcal{Y})=&\,2^{-nT(K+M-1)/2}\pi^{-nT(K+M+1)/2}\sigma_u^{-nT}\sigma_\upsilon^{-nT}\\&\prod_{k=1}^{K}\sigma^{-nT}_{\upsilon^x_k}\prod_{m=1}^{M-1}\sigma^{-nT}_{\upsilon^y_m}\prod_{i=1}^{n}\prod_{t=1}^{T}||\ \mathcal{J}_{it}\left(\theta;\mathcal{Y}_{it}\right)||\\&\prod_{i=1}^{n}\prod_{t=1}^{T}\int_0^\infty\exp\left\{-\frac{1}{2}\left(\frac{\upsilon_{it}\left(\theta,u_{it};\mathcal{Y}_{it}\right)^2}{\sigma^2_\upsilon}+\sum_{k=1}^{K}\frac{\upsilon^x_{k,it}\left(\theta;\mathcal{Y}_{it}\right)^2}{\sigma^2_{\upsilon^x_k}}\right.\right.\\&\left.\left.+\sum_{m=1}^{M-1}\frac{\upsilon^y_{m,it}\left(\theta;\mathcal{Y}_{it}\right)^2}{\sigma^2_{\upsilon^y_m}}+\frac{u^2_{it}}{\sigma^2_u}\right)\right\}du_{it},\end{aligned}\tag{18}$$

where

$$\mathcal{J}_{it}\left(\theta;\mathcal{Y}_{it}\right)=\frac{\partial\mathbf{F}(\theta,u_{it};\mathcal{Y}_{it})}{\partial\left(\ln x_{it},\ln y_{it}\right)}\tag{19}$$

is the Jacobian matrix of the error terms with respect to all logarithmic input and output quantities, which we compute numerically. Notice that the latent u_{it} has to be integrated out of the likelihood function.[7] As the likelihood

[6] Similar stochastic specifications can be found in Malikov et al. (2015) and Tsionas et al. (2015).

[7] While this model specification does not assume that error terms $\mathbf{v}_{it}$ are independent of the input and output quantities, it assumes that the inefficiency term u_{it} is independent of the input and output quantities. This assumption is typical in the literature.

function (18) is complex and depends on latent inefficiency, we use a Bayesian approach with the following prior:

$$p(\theta,\sigma_u,\boldsymbol{\Sigma}) \propto \mathcal{I}_{\mathcal{M}(\mathcal{Y})}(\theta)\sigma_u^{-(\underline{n}+1)}\exp\left\{-\frac{\underline{a}}{2\sigma_u^2}\right\}\sigma_\upsilon^{-(\underline{n}^*+1)}\exp\left\{-\frac{\underline{a}^*}{2\sigma_\upsilon^2}\right\} \tag{20}$$
$$\prod_{k=1}^{K}\sigma_{\upsilon_k^x}^{-(\underline{n}^*+1)}\exp\left\{-\frac{\underline{a}^*}{2\sigma_{\upsilon_k^x}^2}\right\}\prod_{m=1}^{M-1}\sigma_{\upsilon_m^y}^{-(\underline{n}^*+1)}\exp\left\{-\frac{\underline{a}^*}{2\sigma_{\upsilon_m^y}^2}\right\},$$

where $\mathcal{I}_{\mathcal{M}(\mathcal{Y})}(\theta)$ denotes an indicator function that is one if the set of parameters θ satisfies the monotonicity restrictions and zero otherwise and $\underline{n}$, $\underline{a}$, $\underline{n}^*$ and $\underline{a}^*$ are scalars to be set by the analyst. We set $\underline{n}=\underline{n}^*=0$ and $\underline{a}=\underline{a}^*=10^{-5}$, which are non-informative choices (relative to the likelihood). By Bayes' theorem, the posterior distribution is:

$$p(\theta,\sigma_u,\boldsymbol{\Sigma}|\mathcal{Y}) \propto \mathcal{L}(\theta,\sigma_u,\boldsymbol{\Sigma};\mathcal{Y})\cdot p(\theta,\sigma_u,\boldsymbol{\Sigma}). \tag{21}$$

We integrate the posterior analytically with respect to parameters $\sigma_{\upsilon_1^x},...,\sigma_{\upsilon_K^x}$, $\sigma_{\upsilon_1^y},...,\sigma_{\upsilon_{M-1}^y}$ using properties of the Inverted Gamma distribution and we use Metropolis within Gibbs sampling to draw from the conditional posterior distributions $\mathbf{u}|\theta,\sigma_u,\sigma_\upsilon,\mathcal{Y}$, $\sigma_u|\theta,\mathbf{u},\sigma_\upsilon,\mathcal{Y}$, $\sigma_\upsilon|\theta,\mathbf{u},\sigma_u,\mathcal{Y}$ and $\theta|\mathbf{u},\sigma_u,\sigma_\upsilon,\mathcal{Y}$ in the augmented posterior:

$$p\left(\theta,\sigma_u,\sigma_\upsilon,\mathbf{u}\mid\mathcal{Y}\right) \propto \sigma_u^{-(nT+\underline{n}+1)}\exp\left\{-\frac{\underline{a}+\sum_{i=1}^n\sum_{t=1}^T u_{it}^2}{2\sigma_u^2}\right\} \tag{22}$$
$$\prod_{i=1}^n\prod_{t=1}^T \|\, J_{it}\left(\theta;\mathcal{Y}_{it}\right)\|\, \sigma_\upsilon^{-(nT+\underline{n}^*+1)}$$
$$\exp\left\{-\frac{\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}\left(\theta,u_{it};\mathcal{Y}_{it}\right)^2}{2\sigma_\upsilon^2}\right\}$$
$$\prod_{k=1}^{K}\left(\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{k,it}^x\left(\theta;\mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}}$$
$$\prod_{m=1}^{M-1}\left(\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{m,it}^y\left(\theta;\mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}}.$$

For the conditional posterior densities $p\left(\mathbf{u}|\theta,\sigma_u,\sigma_\upsilon,\mathcal{Y}\right)$, $p\left(\sigma_u|\theta,\sigma_\upsilon,\mathbf{u},\mathcal{Y}\right)$ and $p\left(\sigma_\upsilon|\theta,\sigma_u,\mathbf{u},\mathcal{Y}\right)$, we use Gibbs sampling and for $p\left(\theta|\sigma_u,\sigma_\upsilon,\mathbf{u},\mathcal{Y}\right)$ we use the random walk Metropolis-Hastings algorithm. Given the importance of monotonicity in efficiency analysis (Henningsen and Henning 2009), we use rejection sampling to impose the monotonicity conditions at all data points (O'Donnell and Coelli 2005; Terrell 1996).

4 Empirical application

We provide an empirical application to data analysed in Malikov, Kumbhakar, and Tsionas (2015). Our sample is an unbalanced panel of US banks with 395 bank–year observations of the 50 banks with the highest volume of assets contained in the data set. We have the following outputs: consumer loans (y_1), real estate loans (y_2), commercial and industrial loans (y_3), securities (y_4), and off-balance-sheet items (y_5). In order to make our analysis invariant to the units of measurement of the outputs, we normalise them by their arithmetic means. The inputs are labour (number of full-time equivalent employees, x_1), physical capital (x_2), purchased funds (x_3), interest-bearing transaction accounts (x_4), and non-transaction accounts (x_5).[8]

In our application, we do not observe output prices. As we need output prices to calculate the left-hand sides of equations (7) and (8), which are part of the estimated equations (15) and (16), we approximate them by a parametric model of the form:

$$\ln\left(\frac{p_{m,it}}{w_{1,it}}\right) = \eta_m^{(1)} + \eta_m^{(2)}\, t + \tfrac{1}{2}\eta_m^{(3)}\, t^2 \qquad (23)$$
$$\forall\, m = 1, ..., M;\ \forall\, i = 1, ..., n; t = 1, ..., T,$$

in which log output prices relative to the first input price are quadratic functions of the time trend.[9]

Our prior for these parameters is as follows:

$$\eta = [\eta_m^{(1)}, \eta_m^{(2)}, \eta_m^{(3)},\, m = 1, ..., M]' \sim \mathcal{N}_{3M}(0,\, hI). \qquad (24)$$

We set parameter $h = 10$ so that the prior is loose but proper. Based on the price model (23), we replace the unobserved output prices in equations (15) and (16) by:

$$p_{m,it} = w_{1,it}\ \exp\left(\eta_m^{(1)} + \eta_m^{(2)}\, t + \tfrac{1}{2}\eta_m^{(3)}\, t^2\right) \qquad (25)$$
$$\forall\, m = 1, ..., M;\ \forall\, i = 1, ..., n; t = 1, ..., T.$$

Pseudocode of the Metropolis within Gibbs algorithm is as follows:

Pseudocode of the Metropolis within Gibbs algorithm

Step 1. Specify initial values $\theta^{(0)}$, $\eta^{(0)}$, $\sigma_u^{(0)}$, $\sigma_v^{(0)}$ and $\mathbf{u}^{(0)}$.
Step 2. Repeat the following steps for $j = 1, 2, \ldots, R$.

---

[8] Our model specification assumes that the input prices w are exogenous, which is a typical assumption in many empirical analyses, e.g., in analyses based on cost minimisation. However, under certain circumstances, the input prices w could be endogenous, e.g., if differences in input prices between banks reflect heterogeneous inputs rather than 'true' differences in input prices (see, e.g., Quiggin and Bui-Lan 1984).

[9] The use of parametric assumptions to approximate unavailable output prices may introduce approximation errors. The absence of output prices (or even input prices) is common in the literature. In deciding to proceed, we demonstrate that our method is applicable even if output prices are unavailable.

Step 2.1. Sample $u_{it}^{(j)} \mid \theta^{(j-1)}, \sigma_u^{(j-1)}, \sigma_\upsilon^{(j-1)}, \eta^{(j-1)}, \mathcal{Y}_{it}$ from the density of the Half-Normal distribution:

$$N^+\Bigg(-\frac{\sigma_u^{2(j-1)}}{\sigma_\upsilon^{2(j-1)}+\sigma_u^{2(j-1)}} \left(\log \| y_{it} \| - F\left(\log\left(x_{it}\left(\mathcal{Y}_{it}\right)\right), \vartheta\left(\mathcal{Y}_{it}\right), \theta\right)\right), \frac{\sigma_\upsilon^{2(j-1)}\sigma_u^{2(j-1)}}{\sigma_\upsilon^{2(j-1)}+\sigma_u^{2(j-1)}}\Bigg) \forall\, i = 1, 2, \ldots, n; t = 1, 2, \ldots, T.$$

Step 2.2. Sample $\sigma_u^{(j)} \mid \theta^{(j-1)}, \sigma_\upsilon^{(j-1)}, \eta^{(j-1)}, \mathbf{u}^{(j)}, \mathcal{Y}$ from the density of the Inverse Gamma (IG) distribution:

$$IG\left(\frac{nT+\underline{n}}{2}, \frac{\underline{a}+\sum_{i=1}^n\sum_{t=1}^T u_{it}^{2(j)}}{2}\right)$$

Step 2.3. Sample $\sigma_\upsilon^{(j)} \mid \theta^{(j-1)}, \sigma_u^{(j)}, \eta^{(j-1)}, \mathbf{u}^{(j)}, \mathcal{Y}$ from the density of the Inverse Gamma (IG) distribution:

$$IG\left(\frac{nT+\underline{n}^*}{2}, \frac{\underline{a}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}\left(\theta^{(j-1)}, u_{it}^{(j)}; \mathcal{Y}_{it}\right)^2}{2}\right)$$

Step 2.4. Sample $\eta^{(j)} \mid \theta^{(j-1)}, \sigma_u^{(j)}, \sigma_\upsilon^{(j)}, \mathbf{u}^{(j)}, \mathcal{Y}$ from kernel:

$$\exp\left\{-\frac{1}{2h^2}\eta'\eta\right\}\prod_{i=1}^n\prod_{t=1}^T \| J_{it}\left(\theta^{(j-1)}, \eta; \mathcal{Y}_{it}\right) \|$$
$$\prod_{k=1}^K\left(\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}^k\left(\theta^{(j-1)}, \eta; \mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}}$$
$$\prod_{m=1}^{M-1}\left(\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}^m\left(\theta^{(j-1)}, \eta; \mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}}$$

Step 2.5. Sample $\theta^{(j)} \mid \sigma_u^{(j)}, \sigma_\upsilon^{(j)}, \eta^{(j)}, \mathbf{u}^{(j)}, \mathcal{Y}$ applying rejection sampling from kernel:

$$\exp\left\{-\frac{\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}\left(\theta, u_{it}^{(j)}; \mathcal{Y}_{it}\right)^2}{2\sigma_\upsilon^2}\right\}\prod_{i=1}^n\prod_{t=1}^T \| J_{it}\left(\theta, \eta^{(j)}; \mathcal{Y}_{it}\right) \|$$
$$\prod_{k=1}^K\left(\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}^k\left(\theta, \eta^{(j)}; \mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}}$$
$$\prod_{m=1}^{M-1}\left(\underline{\alpha}^*+\sum_{i=1}^n\sum_{t=1}^T \upsilon_{it}^m\left(\theta, \eta^{(j)}; \mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}}$$

There is a potential problem with the present and other similar models in that (i) we arbitrarily selected w_1 as *numeraire* in the price model (23) and (ii) the results depend on the ordering of outputs. We address these concerns by considering different input prices as *numeraire* and all $M!$ possible orderings of outputs as different models. With 5 outputs, there exist $5! = 120$ different orderings of the outputs. As the SRPF is invariant to the ordering of the last two outputs (see Henningsen et al. 2017), there are $5!/2 = 60$ different model specifications regarding the ordering of outputs. To take up the issue of selecting a particular ordering or combining the results from different orderings, we repeat the inference procedure for all 60 different model specifications and we use the marginal likelihood to weight the results obtained from each ordering of outputs using model averaging.

Following Gelfand and Dey (1994) and DiCiccio et al. (1997), the reciprocal of marginal likelihood $\mathcal{M}_c(\mathcal{Y}) = \int_{\omega\in\Omega} p_c(\omega \mid \mathcal{Y})\, d\omega$ for each model $c \in \mathcal{C}$ can be obtained by:

$$\mathcal{M}_c(\mathcal{Y})^{-1} = \mathbb{E}_{p_c(\omega|\mathcal{Y})}\left[\frac{g_c(\omega)}{p_c(\omega \mid \mathcal{Y})} \mid \mathcal{Y}\right], \tag{26}$$

where $\omega = (\theta, \eta, \sigma_u, \sigma_v)'$ and $g(.)$ is a proper density that plays the role of an importance sampling density that closely approximates the posterior distribution:

$$\begin{aligned} p(\omega \mid \mathcal{Y}) \propto \left(\sigma_v^2 + \sigma_u^2\right)^{-\frac{nT}{2}} \sigma_u^{-(\underline{n}+1)} \exp\left\{-\frac{\underline{a}}{2\sigma_u^2}\right\} \exp\left\{-\frac{1}{2h^2}\eta'\eta\right\} \\ \exp\left\{-\frac{1}{2\left(\sigma_v^2+\sigma_u^2\right)} \sum_{i=1}^{n}\sum_{t=1}^{T}\left(\log \| y_{it} \| - F\left(\log\left(x_{it}\right), \theta\left(\mathcal{Y}_{it}\right)\right)\right)^2\right\} \\ \sigma_v^{-(\underline{n}^*+1)} \exp\left\{-\frac{\underline{\alpha}^*}{2\sigma_v^2}\right\} \prod_{i=1}^{n}\prod_{t=1}^{T} \| J_{it}\left(\theta, \eta; \mathcal{Y}_{it}\right) \| \\ \prod_{k=1}^{K}\left(\underline{\alpha}^* + \sum_{i=1}^{n}\sum_{t=1}^{T} \upsilon_{it}^{k}\left(\theta, \eta; \mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}} \\ \prod_{m=1}^{M-1}\left(\underline{\alpha}^* + \sum_{i=1}^{n}\sum_{t=1}^{T} \upsilon_{it}^{m}\left(\theta, \eta; \mathcal{Y}_{it}\right)^2\right)^{-\frac{nT+\underline{n}^*}{2}} \end{aligned} \tag{27}$$

To compute an estimate of the marginal likelihood, we use the following expression:

$$\widehat{\mathcal{M}_c}(\mathcal{Y})^{-1} = \frac{1}{R}\sum_{j=1}^{R} \frac{\phi\left(\omega_c^{(j)}; \widehat{\omega}_c, \widehat{V}_c\right)}{p_c\left(\omega_c^{(j)} \mid \mathcal{Y}\right)} \quad \forall c \in \mathcal{C}, \tag{28}$$

where $\phi(.; m, S)$ denotes the density function of a multivariate normal distribution with mean vector m and variance matrix S and $\widehat{\omega}_c$ and $\widehat{V}_c$ denote the sample average and sample covariance matrix, respectively, computed using the MCMC draws $\left\{\omega_c^{(j)}\right\}_{j=1}^{R}$. Different models are weighted using the posterior model probabilities:

$$\widehat{\pi}_c\left(\mathcal{Y}\right) = \frac{\widehat{\mathcal{M}_c}\left(\mathcal{Y}\right)}{\sum_{c' \in \mathcal{C}} \widehat{\mathcal{M}_{c'}}\left(\mathcal{Y}\right)} \quad \forall c \in \mathcal{C}. \tag{29}$$

For a given technological metric that is a function of the parameters and the data, say $f\left(\omega, \mathcal{Y}\right)$, the related estimate of this metric corresponding to model c is computed as $\widehat{f}_c\left(\mathcal{Y}\right) = \frac{1}{R}\sum_{j=1}^{R} f\left(\omega_c^{(j)}, \mathcal{Y}\right) \; \forall \; c \in \mathcal{C}$ and the model-averaged estimate of this metric is computed as:

$$\sum_{c \in \mathcal{C}} \widehat{\pi}_c\left(\mathcal{Y}\right) \widehat{f}_c\left(\mathcal{Y}\right). \tag{30}$$

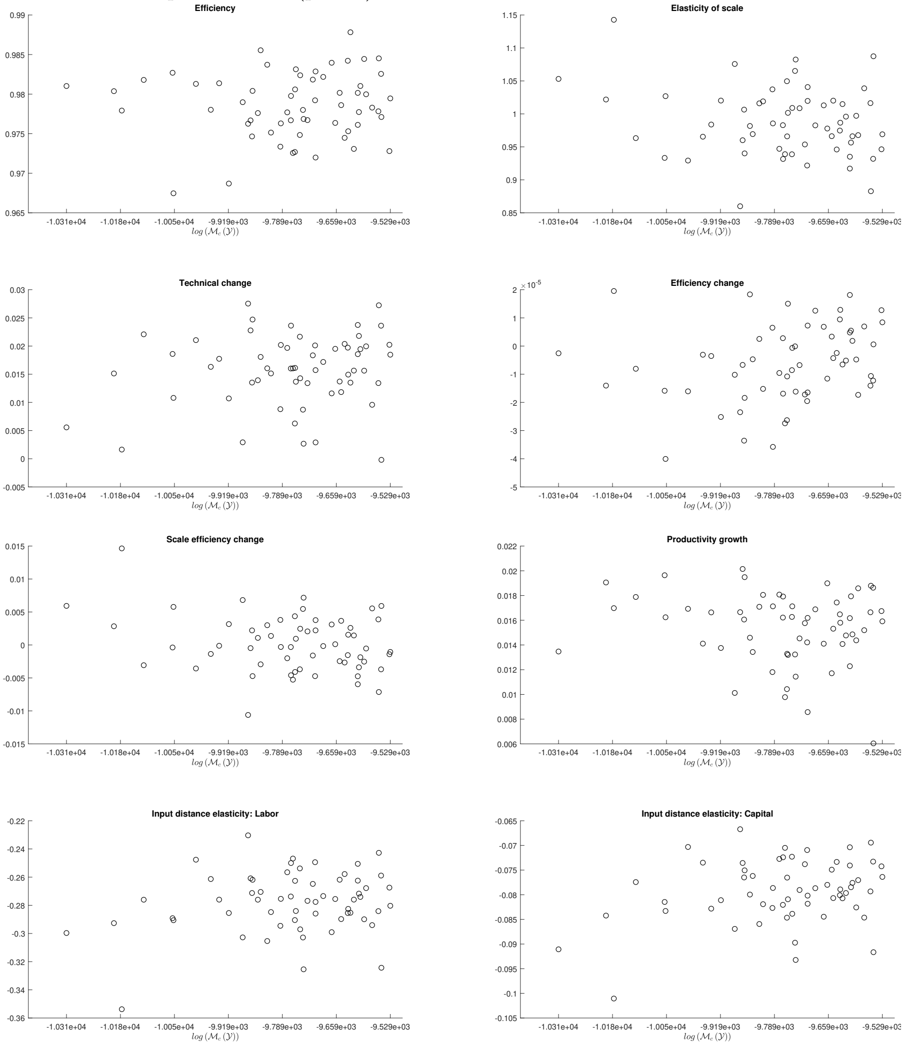
For each of the 60 model specifications, we have repeated the estimation procedure using 25,000 preliminary draws which were discarded followed by another 50,000 draws used to produce estimates of the technological metrics for each model. Despite the high dimensionality, the application of the Metropolis-Hastings algorithm resulted in approximately 30% of all proposals being eventually accepted. Instead of examining the parameter estimates, we focus our attention on more informative technological metrics given by technical efficiency[10], elasticity of scale, technical change, efficiency change, scale efficiency change, and productivity growth. Figures 1 and 2 present scatter plots depicting the average values of the above technological metrics and average values of the distance elasticities against the logarithm of the marginal likelihood for all 60 alternative model specifications.[11] Our results were not sensitive to which input price was used as *numeraire* (up to Monte Carlo errors).

Once the results corresponding to the 60 different output orderings were obtained, model-averaged estimates for the technological metrics were computed as specified in equation (30). Summary statistics of the technological metrics of interest, as obtained through the model averaging procedure, are reported in Table 1. Technical efficiency levels are quite high presenting an average (and median) value of 0.979 with the corresponding 5%–95% interval ranging from 0.976 to 0.982. Elasticities of scale are close to unity presenting a mean value of 0.967, a 5% percentile value of 0.724 and a 95% percentile value of 1.187. The (annual rate of) productivity growth holds most of its mass concentrated around 0.016 with a 5% percentile value of -0.040 and a 95%

[10] To measure observation-specific technical efficiency, we use $\frac{1}{R}\sum_{j=1}^{R} \exp\left(-u_{it}^{(j)}\right) \; \forall \; i = 1, ..., n; t = 1, ..., T$, where $\left\{u_{it}^{(j)}\right\}_{j=1}^{R}$ is a MCMC sample drawn from the posterior.

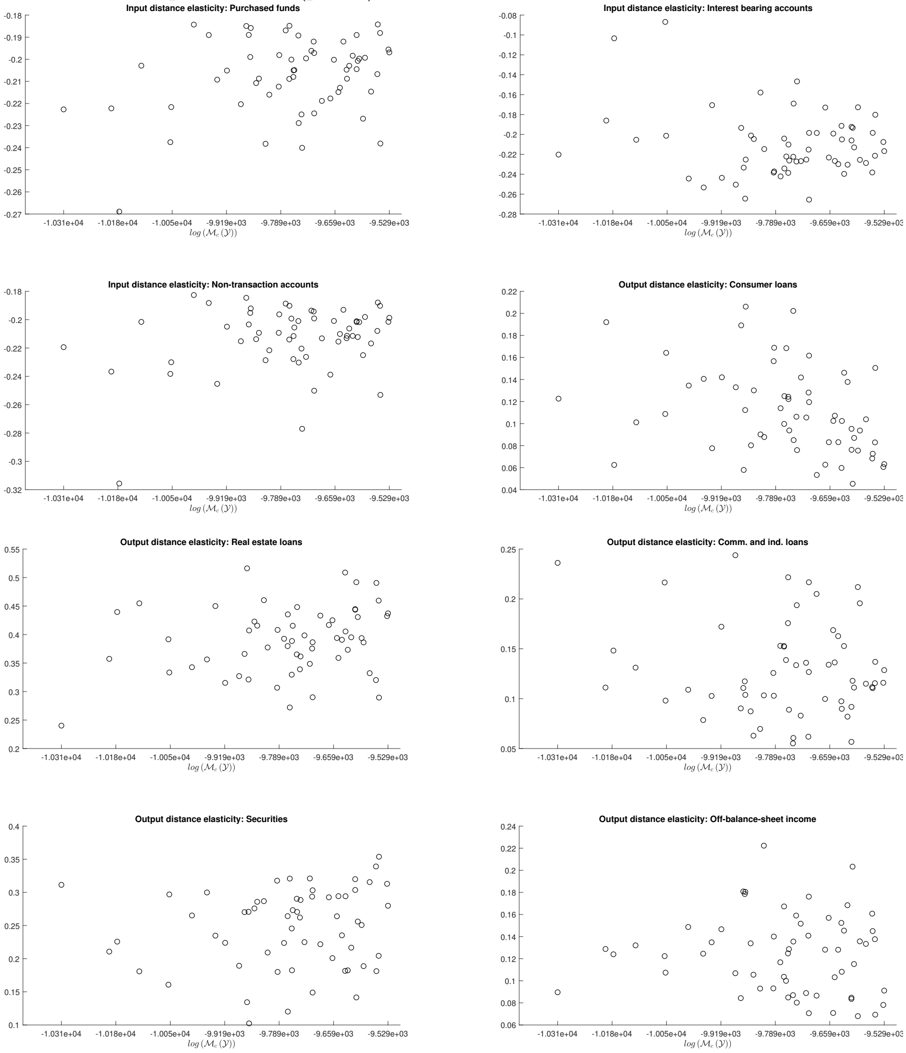
[11] More information on the sensitivity of our results to the ordering of the outputs can be found in the Online Supplement to this Article, where we present detailed statistics of the technological metrics for all 60 different orderings of the outputs.

Fig. 1 Scatter plots between posterior model probabilities and technological metrics for all different model specifications (part 1)



percentile value of 0.065. Technical change presents an average value of 0.019, a 5% percentile value of −0.039 and a 95% percentile value of 0.058, while scale efficiency change averages at −0.001 and its 5%–95% interval ranges from −0.027 to 0.035. Efficiency change averages approximately to zero and its 5% and 95% percentiles are given by −0.001 and 0.001, respectively. As constrained by the imposition of rejection sampling, distance elasticities of inputs (IDE) are negative and those of outputs (ODE) are positive at all data points. In Figure 3, we present kernel density plots that illustrate the distri-

Fig. 2 Scatter plots between posterior model probabilities and technological metrics for all different model specifications (part 2)



butions of model-averaged technological metrics over all observations in our data set.

The results we obtained are in accordance with those reported in other similar efficiency studies of the banking sector. Malikov, Kumbhakar, and Tsionas (2015) use a stochastic directional technology distance function formulation to estimate banking technology in the presence of undesirable outputs, applying Bayesian methods to an extended version of the data set that we use in the present analysis. Their results suggest a mean technical efficiency at the

Table 1 Statistics of model-averaged technological metrics

Technological metrics	Mean	St. Dev.	Median	5% Perc.	95% Perc.
Efficiency	0.9787	0.0016	0.9786	0.9763	0.9816
Elasticity of scale	0.9666	0.1434	0.9776	0.7242	1.1874
Technical change	0.0187	0.0279	0.0225	-0.0391	0.0576
Efficiency change	$9 \cdot 10^{-6}$	0.0007	$-3 \cdot 10^{-5}$	-0.0009	0.0011
Scale efficiency change	-0.0011	0.0704	0.0002	-0.0272	0.0349
Productivity growth	0.0160	0.0706	0.0205	-0.0396	0.0653
IDE: Labour	-0.2790	0.0682	-0.2759	-0.3990	-0.1800
IDE: Capital	-0.0762	0.0199	-0.0763	-0.1082	-0.0432
IDE: Purchased funds	-0.1968	0.0625	-0.2009	-0.2861	-0.0910
IDE: Interest bearing accounts	-0.2158	0.1066	-0.1973	-0.4669	-0.0722
IDE: Non-transaction accounts	-0.1990	0.0739	-0.1957	-0.3317	-0.0740
ODE: Consumer loans	0.0629	0.0424	0.0638	0.0070	0.1082
ODE: Real estate loans	0.4367	0.1801	0.4492	0.1141	0.6997
ODE: Comm. and ind. loans	0.1273	0.0590	0.1203	0.0463	0.2437
ODE: Securities	0.2834	0.1335	0.2723	0.1168	0.5561
ODE: Off-balance-sheet income	0.0897	0.0696	0.0715	0.0143	0.2299

$0.943-0.964$ level, mean desirable scale elasticity at the $0.899-1.021$ level, mean productivity growth at the $0.011-0.018$ level, mean technical change at the $0.011-0.018$ level, and zero mean efficiency change. The technical efficiency results that we present are also in line with those reported in Tsionas (2006), where technical efficiency inferences are obtained for the case of a dynamic stochastic frontier model applied to a panel of large US commercial banks using Gibbs sampling. The mean and median of technical efficiency reported are 0.955 and 0.990, respectively.

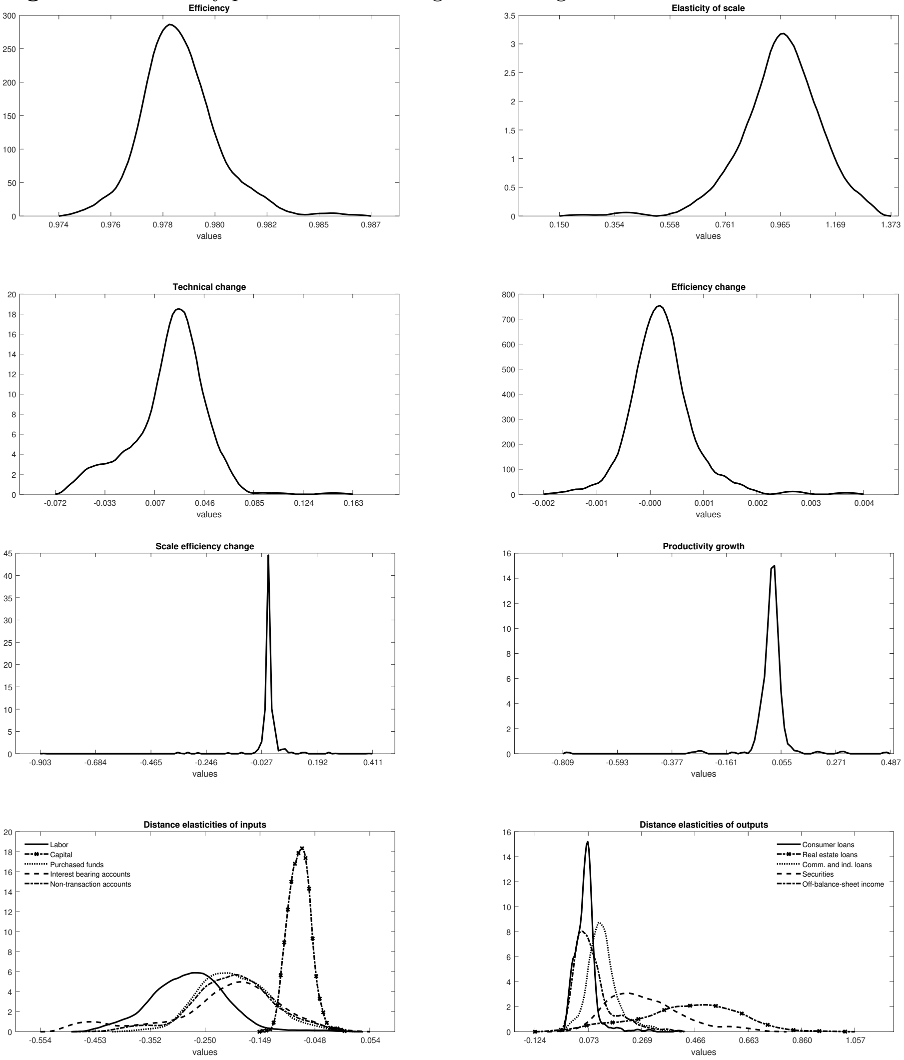
5 Concluding remarks

In this paper, we apply Bayesian methods to provide inference for the SRPF and for technological metrics such as technical inefficiency that avoid possible endogeneity problems which may arise under classical econometric techniques. We address these concerns by estimating the SRPF together with additional equations derived from profit maximization. Despite the simplicity of the functional form, the first-order conditions from profit maximization are complicated and technical inefficiency enters in a non-trivial way throughout the system. Relatively straightforward MCMC techniques have been shown to work well in a substantive application to US banking.

Conflict of interest

All authors declare that they have no conflict of interest.

Fig. 3 Kernel density plots of model-averaged technological metrics



Availability of data and code

The data set and Matlab code are available from the authors upon request and will be made available at the journal's website and/or at the authors' personal websites.

References

Aigner D, Lovell C, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. Journal of Econometrics 6(1):21–37

Akhavein JD, Swamy PAVB, Taubman SB, Singamsetti RN (1997) A general method of deriving the inefficiencies of banks from a profit function. Journal of Productivity Analysis 8(1):71–93

Atkinson SE, Dorfman JH (2005) Bayesian measurement of productivity and efficiency in the presence of undesirable outputs: crediting electric utilities for reducing air pollution. Journal of Econometrics 126(2):445–468

Barten AP (1969) Maximum likelihood estimation of a complete system of demand equations. European Economic Review 1(1):7–73

Battese GE, Coelli TJ (1992) Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis 3(1):153–169, 10.1007/BF00158774

Battese GE, Corra GS (1977) Estimation of a production frontier model: With application to the pastoral zone of eastern Australia. Australian Journal of Agricultural Economics 21(3):169–179

Berger AN, Mester LJ (1997) Inside the black box: What explains differences in the efficiencies of financial institutions? Journal of Banking & Finance 21(7):895–947

Bhattacharyya A, Pal S (2013) Financial reforms and technical efficiency in Indian commercial banking: A generalized stochastic frontier analysis. Review of Financial Economics 22(3):109–117

van den Broeck J, Koop G, Osiewalski J, Steel MF (1994) Stochastic frontier models: A bayesian perspective. Journal of Econometrics 61(2):273–303

Brümmer B, Glauben T, Thijssen G (2002) Decomposition of productivity growth using distance functions: The case of dairy farms in three European countries. American Journal of Agricultural Economics 84(3):628–644

Coelli T, Perelman S (1999) A comparison of parametric and non-parametric distance functions: With application to European railways. European Journal of Operational Research 117(2):326–339

Coelli T, Perelman S (2000) Technical efficiency of European railways: A distance function approach. Applied Economics 32(15):1967–1976

Cuesta RA, Orea L (2002) Mergers and technical efficiency in spanish savings banks: A stochastic distance function approach. Journal of Banking & Finance 26(12):2231–2247

DiCiccio TJ, Kass RE, Raftery A, Wasserman L (1997) Computing Bayes factors by combining simulation and asymptotic approximations. Journal of the American Statistical Association 92:903–915

Färe R, Grosskopf S (1994) Cost and Revenue Constrained Production. Springer

Färe R, Primont D (1995) Multiple-Output Production and Duality: Theory and Applications. Kluwer Academic Publishers

Färe R, Primont D (1996) The opportunity cost of duality. Journal of Productivity Analysis 7(2):213–224

Färe R, Grosskopf S, Lovell CAK, Yaisawarng S (1993) Derivation of shadow prices for undesirable outputs A distance function approach. The Review of Economics and Statistics 75(2):374–380

Ferrier GD, Lovell C (1990) Measuring cost efficiency in banking: Econometric and linear programming evidence. Journal of Econometrics 46(1):229–245

Filippini M, Farsi M (2004) An empirical analysis of cost efficiency in non-profit and public nursing homes. Annals of Public and Cooperative Economics 75:339–365

Gelfand AE, Dey DK (1994) Bayesian model choice: asymptotics and exact calculations. Journal of the Royal Statistical Society B(56):501–514

Gerdtham UG, Löthgren M, Tambour M, Rehnberg C (1999) Internal markets and health care efficiency: a multiple-output stochastic frontier analysis. Health Economics 8(2):151–164

Greene W (2005) Fixed and random effects in stochastic frontier models. Journal of Productivity Analysis 23(1):7–32

Grosskopf S, Hayes KJ, Taylor LL, Weber WL (1997) Budget-constrained frontier measures of fiscal equality and efficiency in schooling. The Review of Economics and Statistics 79(1):116–124

Henningsen A, Henning CHCA (2009) Imposing regional monotonicity on translog stochastic production frontiers with a simple three-step procedure. Journal of Productivity Analysis 32(3):217–229

Henningsen A, Bělín M, Henningsen G (2017) New insights into the stochastic ray production frontier. Economics Letters 156:18–21

Henningsen G, Henningsen A, Jensen U (2015) A Monte Carlo study on multiple output stochastic frontiers: A comparison of two approaches. Journal of Productivity Analysis 44(3):309–320

Huang Th, Wang Mh (2004) Comparisons of economic inefficiency between output and input measures of technical inefficiency using the Fourier flexible cost function. Journal of Productivity Analysis 22(1):123–142

Humphrey DB, Pulley LB (1997) Banks' responses to deregulation: Profits, technology, and efficiency. Journal of Money, Credit and Banking 29(1):73–93

Jondrow J, Lovell] CK, Materov IS, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics 19(2):233–238

Koop G, Steel MF (2001) Bayesian analysis of stochastic frontier models. In: Baltagi BH (ed) A Companion to Theoretical Econometrics, John Wiley & Sons, Ltd, chap 24, pp 520–537

Koop G, Osiewalski J, Steel MFJ (1994) Bayesian efficiency analysis with a flexible form: The aim cost function. Journal of Business & Economic Statistics 12(3):339–346

Koop G, Steel M, Osiewalski J (1995) Posterior analysis of stochastic frontier models using Gibbs sampling. Computational Statistics 10:353–373

Koop G, Osiewalski J, Steel MF (1997) Bayesian efficiency analysis through individual effects: Hospital cost frontiers. Journal of Econometrics 76(1):77–105

Kumbhakar SC (1990) Production frontiers, panel data, and time-varying technical inefficiency. Journal of Econometrics 46(1):201–211, 10.1016/0304-4076(90)90055-X

Kumbhakar SC (2001) Estimation of profit functions when profit is not maximum. American Journal of Agricultural Economics 83(1):1–19

Kumbhakar SC, Bhattacharyya A (1992) Price distortions and resource-use efficiency in Indian agriculture: A restricted profit function approach. The Review of Economics and Statistics 74(2):231–239

Kumbhakar SC, Heshmati A (1995) Efficiency measurement in Swedish dairy farms: An application of rotating panel data, 1976-88. American Journal of Agricultural Economics 77(3):660–674

Kumbhakar SC, Lovell CAK (2000) Stochastic Frontier Analysis. Cambridge University Press, Cambridge

Löthgren M (1997) Generalized stochastic frontier production models. Economics Letters 57:255–259

Löthgren M (2000) Specification and estimation of stochastic multiple-output production and technical inefficiency. Applied Economics 32(12):1533–1540

Lovell CAK, Travers P, Richardson S, Wood L (1994) Resources and functionings: A new view of inequality in Australia. In: Eichhorn W (ed) Models and Measurement of Welfare and Inequality, Springer, pp 787–807, 10.1007/978-3-642-79037-9_41

Malikov E, Kumbhakar SC, Tsionas MG (2015) A cost system approach to the stochastic directional technology distance function with undesirable outputs: The case of US banks in 2001–2010. Journal of Applied Econometrics 31(7):1407–1429

Meeusen W, van den Broeck J (1977) Efficiency estimation from cobb-douglas production functions with composed error. International Economic Review 18(2):435–444

Niquidet K, Nelson H (2010) Sawmill production in the interior of British Columbia: A stochastic ray frontier approach. Journal of Forest Economics 16(4):257–267

O'Donnell CJ, Coelli TJ (2005) A Bayesian approach to imposing curvature on distance functions. Journal of Econometrics 126(2):493–523

Orea L, Kumbhakar SC (2004) Efficiency measurement using a latent class stochastic frontier model. Empirical Economics 29(1):169–183

Pitt MM, Lee LF (1981) The measurement and sources of technical inefficiency in the Indonesian weaving industry. Journal of Development Economics 9(1):43–64

Quiggin J, Bui-Lan A (1984) The use of cross-sectional estimates of profit functions for tests of relative efficiency: A critical review. Australian Journal of Agricultural Economics 28(1):44–55

Rosko MD (2001) Cost efficiency of us hospitals: a stochastic frontier approach. Health Economics 10(6):539–551

Schmidt P, Sickles RC (1984) Production frontiers and panel data. Journal of Business & Economic Statistics 2(4):367–374

Shephard RW (1970) Theory of Cost and Production Functions. Princeton University Press

Sickles RC, Good DH, Getachew L (2002) Specification of distance functions using semi- and nonparametric methods with an application to the dynamic performance of eastern and western European air carriers. Journal of Productivity Analysis 17(1):133–155

Terrell D (1996) Incorporating monotonicity and concavity conditions in flexible functional forms. Journal of Applied Econometrics 11:179–194

Tsionas EG (2000) Full likelihood inference in normal-gamma stochastic frontier models. Journal of Productivity Analysis 13(3):183–205

Tsionas EG (2002) Stochastic frontier models with random coefficients. Journal of Applied Econometrics 17(2):127–147

Tsionas EG (2006) Inference in dynamic stochastic frontier models. Journal of Applied Econometrics 21(5):669–676

Tsionas EG, Kumbhakar SC, Malikov E (2015) Estimation of input distance functions: A system approach. American Journal of Agricultural Economics 97(5):1478–1493

Vivas AL (1997) Profit efficiency for Spanish savings banks. European Journal of Operational Research 98(2):381–394