DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

# IFRO Working Paper

# Estimating Stochastic Ray Production Frontiers

*Mike Tsionas*
*Marwan Izzeldin*
*Arne Henningsen*
*Evaggelos Paravalos*

**2019 / 06**

**IFRO Working Paper 2019 / 06**

Estimating Stochastic Ray Production Frontiers

Authors: Mike Tsionas, Marwan Izzeldin, Arne Henningsen, Evaggelos Paravalos

JEL-classification: C11, C13, D24

Published: September 2019

See the full series IFRO Working Paper here:
www.ifro.ku.dk/english/publications/ifro_series/working_papers/

Department of Food and Resource Economics (IFRO)
University of Copenhagen
Rolighedsvej 25
DK 1958 Frederiksberg  DENMARK
www.ifro.ku.dk/english/

# Estimating Stochastic Ray Production Frontiers

**Mike Tsionas[a], Marwan Izzeldin[a], Arne Henningsen[b*], and Evaggelos Paravalos[c]**

[a] Department of Economics, Lancaster University Management School, LA1 4YX UK

[b] Department of Food and Resource Economics, University of Copenhagen, Rolighedsvej 25, 1958 Frederiksberg C, Denmark

[c] Department of Economics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece

[*] Corresponding author, e-mail: arne@ifro.ku.dk

September 9, 2019

**Abstract**

In this paper, we consider the stochastic ray production function that has been revived recently by Henningsen et al. (2017). We use a profit-maximizing framework to resolve endogeneity problems that are likely to arise, as in all distance functions, and we derive the system of equations after incorporating technical inefficiency. As technical inefficiency enters non-trivially into the system of equations and the Jacobian is highly complicated, we propose Monte Carlo methods of inference. We illustrate the new approach using US banking data and we also address the problems of missing prices and selection of ordering for outputs.

**Keywords:** Stochastic ray production frontier, Technical inefficiency, Profit maximization, Bayesian inference

**JEL codes:** C11, C13, D24

## 1 Introduction

Henningsen et al. (2017) have reinstated the stochastic ray production frontier (SRPF) proposed by Löthgren (1997). The SRPF 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 an output distance function (ODF), i.e., $D(x, y) = \min\left\{\lambda > 0 :\ (y\lambda^{-1}, x) \text{ can be produced}\right\}$ with $0 < D(x, y) \leq 1$, can be transformed to a SPRF specified as:

$$\ln ||y|| = F\left(\ln x, \vartheta(y)\right) + v - u, \tag{1}$$

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

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

which can be estimated as a stochastic frontier model. However, the econometric estimation of equation (1) likely gives inconsistent estimates, because the explanatory variables $x$ (input quantities) and $\vartheta(y)$ (angles between the output quantities) are endogenous decision variables and, thus, are likely correlated with the inefficiency term $u$ and/or the noise term $v$. In this paper, we suggest an econometric procedure for the consistent estimation of the SRPF.

## 2 Microeconomic specification and first-order conditions

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 take into account the endogeneity of $x$ and $\vartheta(y)$ by estimating an SRPF (2) along with the corresponding profit-maximizing FOC. We derive the FOC from the profit maximization problem:

$$\max_{x \in \Re^{+}_{K}, y \in \Re^{+}_{M}} p'y - w'x, \ \text{s.t.}\ D(x, y) = e^{-u}, \tag{3}$$

[1] 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 $\lambda$ denotes the Lagrange multiplier, the FOC are:

$$w_k = -\lambda \frac{\partial D(x,y)}{\partial x_k} = -\lambda \frac{\partial \ln D(x,y)}{\partial \ln x_k} \frac{D(x,y)}{x_k} \; \forall \, k = 1, ..., K, \tag{4}$$

$$p_m = \lambda \frac{\partial D(x,y)}{\partial y_m} = \lambda \frac{\partial \ln D(x,y)}{\partial \ln y_m} \frac{D(x,y)}{y_m} \; \forall \, m = 1, ..., M, \tag{5}$$

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

$$\frac{w_k x_k}{p'y} = -\frac{\partial \ln D(x,y)}{\partial \ln x_k} \; \forall \, k = 1, ..., K, \tag{6}$$

$$\frac{p_m y_m}{p'y} = \frac{\partial \ln D(x,y)}{\partial \ln y_m} \; \forall \, m = 1, ..., M. \tag{7}$$

The right-hand sides of equations (6) and (7) are the distance elasticities of the inputs and outputs, respectively, that can be calculated as:[2]

$$\frac{\partial \ln D(\ln x, \vartheta)}{\partial \ln x_k} = -\frac{\partial F(\ln x, \vartheta)}{\partial \ln x_k} = -\beta_k - \sum_{k'=1}^{K} \beta_{kk'} \ln x_{k'} - \sum_{m=1}^{M-1} \gamma_{mk} \vartheta_m - \zeta_{\mathfrak{T}k}^x t \; \forall \, k = 1, ..., K, \tag{8}$$

$$\frac{\partial \ln D(\ln x, \vartheta)}{\partial \ln y_m} = \frac{\partial \ln ||y||}{\partial \ln y_m} - \frac{\partial F(\ln x, \vartheta)}{\partial \ln y_m} = \frac{y_m^2}{||y||^2} - \sum_{m'=1}^{\min(m,M-1)} \frac{\partial F(\ln x, \vartheta)}{\partial \vartheta_{m'}} \cdot \frac{\partial \vartheta_{m'}}{\partial \ln y_m} \tag{9}$$

$$= \frac{y_m^2}{||y||^2} + \sum_{m'=1}^{\min(m,M-1)} \left( \alpha_{m'} + \sum_{m^*=1}^{M-1} \alpha_{m'm^*} \vartheta_{m^*} + \sum_{k=1}^{K} \gamma_{m'k} \ln x_k + \zeta_{\mathfrak{T}m'}^y t \right) \tag{10}$$

$$\frac{y_m S_{m'} \left( \delta_{mm'} - y_m y_{m'} S_{m'}^2 \right)}{\sqrt{1 - y_{m'}^2 S_{m'}^2}} \; \forall \, m = 1, ..., M,$$

where $S_{m'} = 1 \Big/ \sqrt{\sum_{m^*=m'}^{M} y_{m^*}^2}$ 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 $\partial \ln D(x,y)/\partial \ln y_m \geq 0 \; \forall \, m = 1, ..., M$ (see Kumbhakar and Lovell, 2000, p. 32).

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

[2]Note that $\ln D(x,y) = -u = \ln ||y|| - F(\ln x, \vartheta(y)) - v$, $\partial \vartheta_{m'} / \partial \ln y_m = (\partial \vartheta_{m'} / \partial y_m) \, y_m$, $\partial \vartheta_{m'} / \partial \ln y_m = 0 \; \forall \, m < m'$, and $\partial \arccos(z) / \partial z = -1/\sqrt{1 - z^2}$.

## 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 (1, 6, 7). 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., $\sum_{m=1}^{M} (p_m y_m) / (p'y) = \sum_{m=1}^{M} \partial \ln D(x,y)$ $/\partial \ln y_m = 1$, one of the FOC regarding the outputs (7) 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 (1), $K$ FOC with respect to the input quantities (6), and $(M-1)$ FOC with respect to the output quantities (7).

Provided we have panel data ($i = 1, ..., n,\ t = 1, ..., T$) and adding error terms to the FOC, the system of equations used in the estimation is:

$$\begin{aligned} v_{it} =& \ln ||y_{it}|| - \alpha_0 - \sum_{m=1}^{M-1} \alpha_m \vartheta_{m,it} - \tfrac{1}{2} \sum_{m=1}^{M-1} \sum_{m'=1}^{M-1} \alpha_{mm'} \vartheta_{m,it} \vartheta_{m',it} \\ & - \sum_{k=1}^{K} \beta_k \ln x_{k,it} - \tfrac{1}{2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \beta_{kk'} \ln x_{k,it} \ln x_{k',it} - \sum_{m=1}^{M-1} \sum_{k=1}^{K} \gamma_{mk} \ln x_{k,it} \vartheta_{m,it} \\ & - \zeta_{\mathfrak{T}} t + \tfrac{1}{2} \zeta_{\mathfrak{T}\mathfrak{T}} \mathbf{t}^2 - \sum_{m=1}^{M-1} \zeta_{\mathfrak{T}m}^{y} t \, \vartheta_{m,it} - \sum_{k=1}^{K} \zeta_{\mathfrak{T}k}^{x} t \, \ln x_{k,it} + u_{it}, \end{aligned} \tag{11}$$

$$v_{k,it}^{x} = \frac{w_{k,it} x_{k,it}}{p'_{it} y_{it}} - \beta_k - \sum_{k'=1}^{K} \beta_{kk'} \ln x_{k'} - \sum_{m=1}^{M-1} \gamma_{mk} \vartheta_m - \zeta_{\mathfrak{T}k}^{x} t \; \forall \, k = 1, ..., K, \tag{12}$$

$$\begin{aligned} v_{m,it}^{y} =& \frac{p_{m,it} y_{m,it}}{p'_{it} y_{it}} - \frac{y_{m,it}^2}{||y_{it}||^2} - \sum_{m'=1}^{\min(m,M-1)} \left( \alpha_{m'} + \sum_{m^*=1}^{M-1} \alpha_{m'm^*} \vartheta_{m^*,it} + \sum_{k=1}^{K} \gamma_{m'k} \ln x_{k,it} + \zeta_{\mathfrak{T}m'}^{y} t \right) \\ & \left( \frac{y_{m,it} S_{m',it} \left( \delta_{mm'} - y_{m,it} y_{m',it} S_{m',it}^2 \right)}{\sqrt{1 - y_{m',it}^2 S_{m',it}^2}} \right) \; \forall \; m = 1, ..., M-1. \end{aligned} \tag{13}$$

We denote the vector of error terms by $\mathbf{v}_{it} \equiv \left[ v_{it}, v_{1,it}^{x}, ..., v_{K,it}^{x}, v_{1,it}^{y}, \ldots, v_{M-1,it}^{y} \right]'$ and assume that it follows a $(K+M)$-variate normal distribution, i.e., $\mathbf{v}_{it} \sim \mathcal{N}_{K+M} (\mathbf{0}, \mathbf{\Sigma}) \; \forall i = 1, ..., n; t = 1, ..., T$, where $\mathbf{\Sigma} = \text{diag}(\sigma_{v}^2, \sigma_{v_1^x}^2, ..., \sigma_{v_K^x}^2, \sigma_{v_1^y}^2, ..., \sigma_{v_{M-1}^y}^2)$. For technical inefficiency we make the standard assumption that: $u_{it} \sim \mathcal{N}^{+} (0, \sigma_u^2)$ 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:

$$\mathbf{v}_{it} = \mathbf{F}(\theta, u_{it}; \mathcal{Y}_{it}), \tag{14}$$

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 $(i, t)$.

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

$$\mathcal{L}(\theta, \sigma_u, \mathbf{\Sigma}; \mathcal{Y}) = 2^{-nT(K+M-1)/2} \pi^{-nT(K+M+1)/2} \sigma_u^{-nT} \sigma_\upsilon^{-nT} \prod_{k=1}^{K} \sigma_{\upsilon_k^x}^{-nT} \prod_{m=1}^{M-1} \sigma_{\upsilon_m^y}^{-nT} \prod_{i=1}^{n} \prod_{t=1}^{T} || \, J_{it} \left(\theta; \mathcal{Y}_{it}\right) || \tag{15}$$

$$\prod_{i=1}^{n} \prod_{t=1}^{T} \int_0^\infty \exp \left\{ -\frac{1}{2} \left( \frac{\upsilon_{it} \left(\theta, u_{it}; \mathcal{Y}_{it}\right)^2}{\sigma_\upsilon^2} + \sum_{k=1}^{K} \frac{\upsilon_{k,it}^x \left(\theta; \mathcal{Y}_{it}\right)^2}{\sigma_{\upsilon_k^x}^2} + \sum_{m=1}^{M-1} \frac{\upsilon_{m,it}^y \left(\theta; \mathcal{Y}_{it}\right)^2}{\sigma_{\upsilon_m^y}^2} + \frac{u_{it}^2}{\sigma_u^2} \right) \right\} du_{it}$$

where

$$\mathcal{J}_{it} \left(\theta; \mathcal{Y}_{it}\right) = \frac{\partial \mathbf{F}(\theta, u_{it}; \mathcal{Y}_{it})}{\partial \left(\ln x_{it}, \ln y_{it}\right)} \tag{16}$$

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. As the likelihood function (15) is complex and depends on latent inefficiency, we use a Bayesian approach with the following prior:

$$p(\theta, \sigma_u, \mathbf{\Sigma}) \propto \mathcal{I}_{\mathcal{M}(\mathcal{Y})}(\theta) \sigma_u^{-(\underline{n}+1)} \exp \left\{ -\frac{\underline{a}}{2\sigma_u^2} \right\} \sigma_\upsilon^{-(\underline{n}^*+1)} \exp \left\{ -\frac{\underline{a}^*}{2\sigma_\upsilon^2} \right\} \prod_{k=1}^{K} \sigma_{\upsilon_k^x}^{-(\underline{n}^*+1)} \exp \left\{ -\frac{\underline{a}^*}{2\sigma_{\upsilon_k^x}^2} \right\} \tag{17}$$

$$\prod_{m=1}^{M-1} \sigma_{\upsilon_m^y}^{-(\underline{n}^*+1)} \exp \left\{ -\frac{\underline{a}^*}{2\sigma_{\upsilon_m^y}^2} \right\},$$

where $\mathcal{I}_{\mathcal{M}(\mathcal{Y})}(\theta)$ denotes an indicator function that is one if the set of parameters $\theta$ 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:

$$p(\theta, \sigma_u, \mathbf{\Sigma} | \mathcal{Y}) \propto \mathcal{L}(\theta, \sigma_u, \mathbf{\Sigma}; \mathcal{Y}) \cdot p(\theta, \sigma_u, \mathbf{\Sigma}). \tag{18}$$

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 $\mathbf{u}|\theta, \sigma_u, \sigma_\upsilon, \mathcal{Y}$, $\theta|\mathbf{u}, \sigma_u, \sigma_\upsilon, \mathcal{Y}$, $\sigma_u|\theta, \mathbf{u}, \sigma_\upsilon, \mathcal{Y}$ and $\sigma_\upsilon|\theta, \mathbf{u}, \sigma_u, \mathcal{Y}$ in the augmented posterior:

$$p\left(\theta, \sigma_u, \sigma_\upsilon, \mathbf{u} \mid \mathcal{Y}\right) \propto \sigma_u^{-(nT+\underline{n}+1)} \exp \left\{ -\frac{\underline{a} + \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2}{2\sigma_u^2} \right\} \prod_{i=1}^{n} \prod_{t=1}^{T} || \, J_{it} \left(\theta; \mathcal{Y}_{it}\right) || \tag{19}$$

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

$$\prod_{k=1}^{K} \left( \underline{\alpha}^* + \sum_{i=1}^{n} \sum_{t=1}^{T} \upsilon_{k,it}^x \left(\theta; \mathcal{Y}_{it}\right)^2 \right)^{-\frac{nT+\underline{n}^*}{2}} \prod_{m=1}^{M-1} \left( \underline{\alpha}^* + \sum_{i=1}^{n} \sum_{t=1}^{T} \upsilon_{m,it}^y \left(\theta; \mathcal{Y}_{it}\right)^2 \right)^{-\frac{nT+\underline{n}^*}{2}}$$

For the conditional posterior densities $p(\mathbf{u}|\theta, \sigma_u, \sigma_v, \mathcal{Y})$, $p(\sigma_u|\theta, \sigma_v, \mathbf{u}, \mathcal{Y})$ and $p(\sigma_v|\theta, \sigma_u, \mathbf{u}, \mathcal{Y})$ we use Gibbs sampling and for $p(\theta|\sigma_u, \sigma_v, \mathbf{u}, \mathcal{Y})$ we use the random walk Metropolis-Hastings algorithm.[3] 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 et al. (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 dataset. 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 ($y_5$). 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 ($x_5$). It should be noted that differences in input prices between banks may reflect heterogeneous inputs rather than 'true' differences in input prices (see, e.g., Quiggin and Bui-Lan, 1984).

In our application, we do not observe output prices. As we need output prices to calculate the left-hand sides of equations (6) and (7), which are part of the estimated equations (12) and (13), we approximate them by a parametric model of the form:

$$\ln\left(\frac{p_{m,it}}{w_{1,it}}\right) = \eta_m^{(1)} + \eta_m^{(2)}\, t + \tfrac{1}{2}\eta_m^{(3)}\, t^2 \;\forall\, m = 1, ..., M;\; \forall\, i = 1, ..., n; t = 1, ..., T, \qquad (20)$$

in which log output prices relative to the first input price are quadratic functions of the time trend. Our prior for these parameters is as follows:

$$\eta = [\eta_m^{(1)}, \eta_m^{(2)}, \eta_m^{(3)},\ m = 1, ..., M]' \sim \mathcal{N}_{3M}(0,\ hI). \qquad (21)$$

We set parameter $h = 10$ so that the prior is loose but proper. Based on the price model (20), we replace the unobserved output prices in equations (12) and (13) by:

$$p_{m,it} = w_{1,it}\ \exp\left(\eta_m^{(1)} + \eta_m^{(2)}\, t + \tfrac{1}{2}\eta_m^{(3)}\, t^2\right) \;\forall\, m = 1, ..., M;\; \forall\, i = 1, ..., n; t = 1, ..., T. \qquad (22)$$

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 (20) 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. Our results were not sensitive to which input price was used as *numeraire* (up to Monte Carlo errors). However, we need to take up the issue of selecting a particular ordering or combining the results from different orderings. We repeated the inference procedure for all 60 different model specifications and we used

[3] Pseudocode describing the Metropolis within Gibbs algorithm applied can be found in the Appendix.

marginal likelihood to weigh the results obtained from each ordering of outputs using model averaging. To estimate marginal likelihood, we followed Chib (1995, formula (5)) and DiCiccio et al. (1997). We have used 25,000 preliminary iterations which have been discarded followed by another 50,000 to obtain our main results. Despite the high dimensionality, the application of the Metropolis-Hastings algorithm resulted in approximately 30% of all proposals being eventually accepted.

Using model averaging, we obtained measures of input and output elasticities, elasticity of scale, technical change, efficiency change, scale efficiency change, productivity growth and inefficiency scores. Although estimated measures of interest present small differences for different orderings, qualitatively they turned out very similar. We report our most important findings in Figure 1. Technical inefficiency averages approximately to 0.018 and ranges from 0.013 to 0.026. Elasticities of scale are close to unity presenting a mean value of 0.933, a minimum value of 0.732 and a maximum value of 1.164. The (annual rate of) productivity growth was estimated between $-0.093$ and 0.284 with most of the mass concentrated around 0.020. Distance elasticities of inputs (outputs) are negative (positive) at all data points, since this has been imposed by rejection sampling.

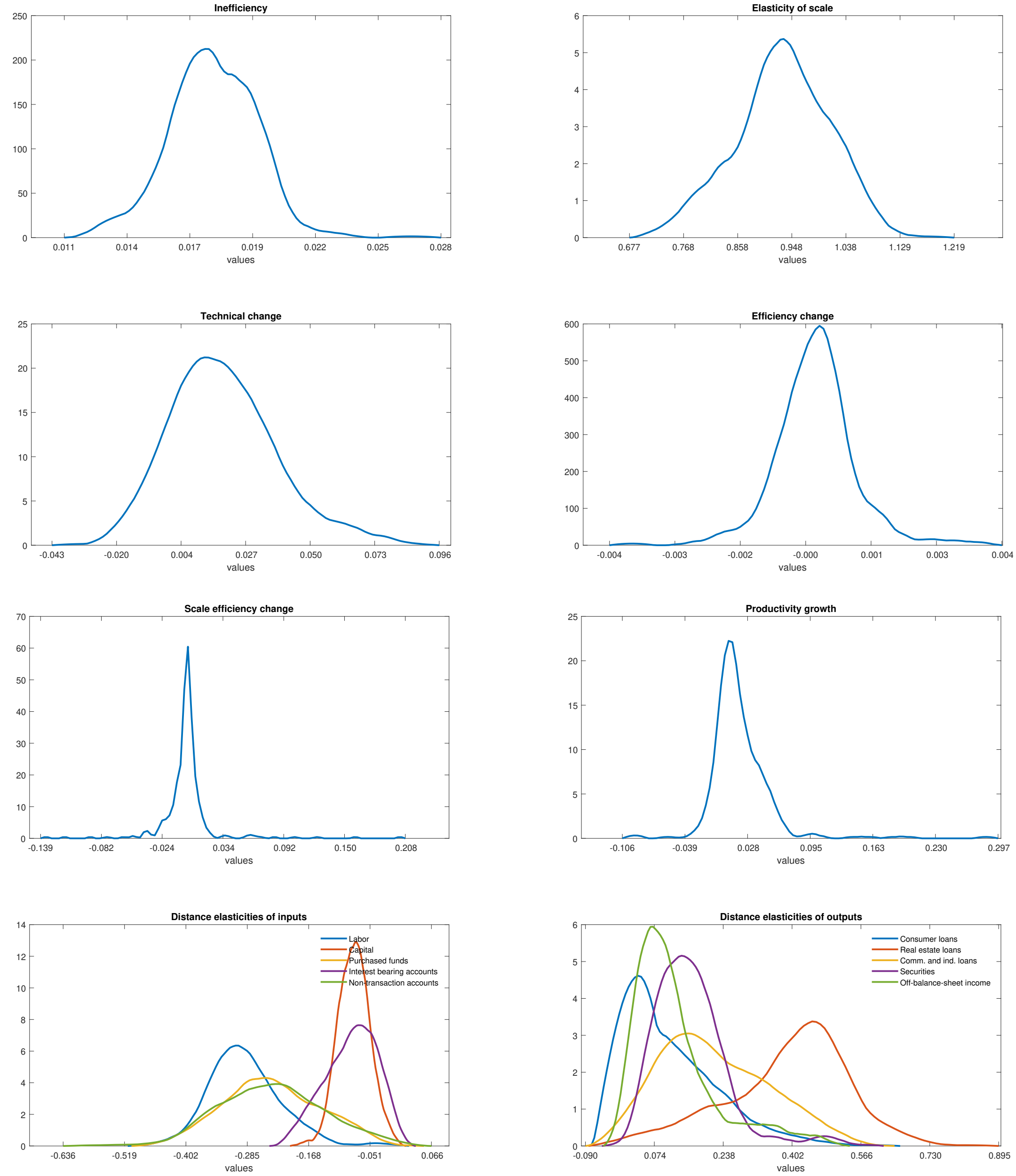
## 5 Concluding remarks

In this paper we have considered the stochastic ray production frontier (SRPF) recently analyzed by Henningsen et al. (2017). We argued that endogeneity problems might be present in the formulation and, therefore, OLS or ML estimation of the frontier might not deliver consistent results. Therefore, one needs to complete the system with additional equations for the endogenous inputs and outputs. 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 U.S. banking.

## Acknowledgements

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

Figure 1: Summaries of the model



## References


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

Brümmer, B., Glauben, T., and 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.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90:1313–1321.

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

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

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

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

Henningsen, G., Henningsen, A., and 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.

Kumbhakar, S. C. and Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge.

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

Malikov, E., Kumbhakar, S. C., and Tsionas, M. G. (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.

O'Donnell, C. J. and Coelli, T. J. (2005). A Bayesian approach to imposing curvature on distance functions. *Journal of Econometrics*, 126(2):493–523.

Quiggin, J. and 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.

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

Tsionas, E. G., Kumbhakar, S. C., and Malikov, E. (2015). Estimation of input distance functions: A system approach. *American Journal of Agricultural Economics*, 97(5):1478–1493.

## Appendix

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 $j = 1, 2, \ldots, R$.

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

$$N^{+}\left(-\frac{\sigma_u^{2(j-1)}}{\sigma_v^{2(j-1)}+\sigma_u^{2(j-1)}}\left(\log \| y_{it} \| - F\left(\log\left(x_{it}\left(\mathcal{Y}_{it}\right)\right), \vartheta\left(\mathcal{Y}_{it}\right), \theta\right)\right), \frac{\sigma_v^{2(j-1)}\sigma_u^{2(j-1)}}{\sigma_v^{2(j-1)}+\sigma_u^{2(j-1)}}\right)$$

$$\forall\, i = 1, 2, \ldots, n; t = 1, 2, \ldots, T.$$

**Step 2.2.** Sample $\sigma_u^{(j)} \mid \theta^{(j-1)}, \sigma_v^{(j-1)}, \eta^{(j-1)}, \mathbf{u}^{(j)}, \mathcal{Y}$ 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 $\sigma_v^{(j)} \mid \theta^{(j-1)}, \sigma_u^{(j)}, \eta^{(j-1)}, \mathbf{u}^{(j)}, \mathcal{Y}$ 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 $\eta^{(j)} \mid \theta^{(j-1)}, \sigma_u^{(j)}, \sigma_v^{(j)}, \mathbf{u}^{(j)}, \mathcal{Y}$ 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 $\theta^{(j)} \mid \sigma_u^{(j)}, \sigma_v^{(j)}, \eta^{(j)}, \mathbf{u}^{(j)}, \mathcal{Y}$ from kernel:

$$\exp\left\{-\frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\upsilon_{it}\left(\theta, u_{it}^{(j)}; \mathcal{Y}_{it}\right)^{2}}{2\sigma_1^2}\right\}\prod_{i=1}^{n}\prod_{t=1}^{T} \| J_{it}\left(\theta, \eta^{(j)}; \mathcal{Y}_{it}\right) \|$$

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