A Spatial Stochastic Frontier Model with Endogenous Frontier and Environmental Variables

Levent Kutlu, Kien C. Tran, and Mike G. Tsionas

Keywords: Banking; Endogeneity; Spatial autoregressive spillover; Stochastic frontier.

JEL Classification: C13, C23, C36, D24.

1. Introduction

In this study, we solve a variety of endogeneity problems, for the first time, for spatial stochastic frontier models. Traditional stochastic frontier models do not control for spatial lag of the dependent variable, which captures so called spatial autoregressive (SAR) dependence (see Cliff and Ord, 1973, 1981). If such a dependence is present, omitting the SAR term would lead to inconsistent parameter and efficiency estimates. Druska and Horrace (2004), Glass, Kenjegalieva, and Paez-Farrell (2013) (GKP), Glass, Kenjegalieva, and Sickles (2014) (GKS), and Kutlu and Nair-Reichert (2018a) solve this problem via distribution-free approaches.[1] An important advantage of these distribution-free approaches is that we do not assume a specific distribution to the inefficiency term. However, outliers may have serious implications for the magnitudes of the efficiency estimates.[2] Hence, alternatively, in the conventional stochastic frontier literature, it is common to represent inefficiency via a one-sided error term. In the spatial spillover context, Glass, Kenjegalieva, and Sickles (2016) follow this approach and introduce the SAR variable while also making distributional assumptions (i.e., half normal distribution) on the inefficiency component of the error structure.[3]

While these stochastic frontier approaches solve the endogeneity problem due to the SAR variable being endogenous, they don't solve the endogeneity problems resulting from the endogeneity of frontier variables (other than SAR term) and environmental variables (i.e., variables that affect inefficiency), which would lead to inconsistent parameter and efficiency estimates. For example, Mutter et al. (2013) argue that if the quality is a part of the production process where it is cost enhancing and quantity and quality decisions are made simultaneously, then the quality variable would be endogenous, i.e., correlated with the two-sided error term.[4] In the stochastic frontier context, there is a recent yet growing interest for solutions to this type of endogeneity problem. Guan et al. (2009), Kutlu (2010),

---

[1] See Schmidt and Sickles (1984) and Cornwell, Schmidt, and Sickles (1990) for non-spatial distribution-free stochastic frontier models and Duygun, Kutlu, and Sickles (2016) for their Kalman filter counterparts.
[2] See Kutlu (2012, 2017) for a more details about this issue and some potential solutions.
[3] See Han, Ryu, and Sickles (2016) for an extension of Glass, Kenjegalieva, and Sickles (2014) where the spatial weighting matrix is time-varying.
[4] Note that dropping the quality variable does not solve the problem as this would bias efficiency estimates.

Tran and Tsionas (2013), Amsler, Prokhorov, and Schmidt (2016, 2017), Griffiths and Hajargasht (2016), Karakaplan and Kutlu (2017a,b) [5], and Kutlu, Tran, and Tsionas (2018) exemplify such studies.[6]

In this study, we consider a SAR stochastic frontier model where endogeneity of both frontier and environmental variables are allowed. Hence, we solve three different endogeneity problems (endogeneity of SAR term, frontier variables, and environmental variables) at the same time. We achieve this by employing a control function approach, which was first introduced by Kutlu (2010) to the stochastic frontier literature. Our general estimation strategy can easily be modified and applied in both cross sectional SAR stochastic frontier context as well as conventional SAR models without inefficiency, i.e., full efficiency. Moreover, besides cost and production function estimation, our model can may be applied in the industrial organization setting where the one-sided error term captures the market power. For example, Orea and Steinbucks (2018), Karakaplan and Kutlu (2018a), and Kutlu and Wang (2018) propose conduct parameter models that use stochastic frontier models. Since the conduct parameter models involve estimation of demand and supply equations, they would suffer from endogeneity issues; and thus these studies utilize techniques that solve endogeneity issues. Therefore, the scope of our contribution is beyond the efficiency measurement context.

2. The Model and Estimation of Efficiency

2.1. The Model

For the sake of fixing the ideas, we present a production function. The same equations can be used for the cost function estimation with minor modifications. We call a variable endogenous if it is

[5] See Karakaplan and Kutlu (2018b) and Kutlu and Nair-Reichert (2018b) for applications of Karakaplan and Kutlu (2017a,b).
[6] Kutlu and Sickles (2012) uses similar approaches to solve endogeneity issues in the Kalman filter estimation context.

correlated with the two-sided error term. Consider the following stochastic frontier model:

$$
\begin{aligned}
y_i &= \rho \sum\nolimits_j w_{ij} y_j + x_{1i}' \beta - u_i + v_i \\
x_i &= z_i \delta + \varepsilon_i \\
u_i &= h_i u_i^* \\
h_i &= f(x_{2i}' \varphi_u) > 0,
\end{aligned}
\tag{1}
$$

where y_i is the logarithm of the output for i^{th} productive unit; w_{ij} is the weight of j^{th} productive unit's output for i^{th} productive unit where $\sum\nolimits_j w_{ij} = 1$, i.e., the weights are so called row-normalized; x_{1i} is a vector of variables that may include endogenous variables; $u_i \geq 0$ is a one-sided term that is capturing the inefficiency; $u_i^* \sim N^+(\mu, \sigma_u^2)$ and x_{2i} is a vector of variables that may include endogenous variables, which does not contain the constant; v_i is the usual two-sided error term for the production function; x_i is a vector of endogenous variables from x_{1i} and/or x_{2i}, i.e., $x_i \subseteq (x_{1i} \cup x_{2i})$; z_i is a matrix of instrumental variables (first row for the first endogenous variable from x_i, second row for second endogenous variable, etc.); ε_i is a vector of usual error terms; and β, δ, and φ_u are parameters. We assume that u_i^* is independent of x_{1i}, x_{2i}, and v_i. Let Ω_ε be the variance-covariance matrix of ε_i, and $\zeta_i^* = (\varepsilon_i^{*\prime}, v_i)' = (\varepsilon_i' \Omega_\varepsilon^{-1/2}, v_i)'$. Also, assume that:

$$
\zeta_i^* \sim \boldsymbol{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_q & \sigma_v \tau \\ \sigma_v \tau' & \sigma_v^2 \end{bmatrix} \right) = \boldsymbol{N}(0, \Omega),
$$

where $\boldsymbol{N}(.,.)$ denotes the multivariate normal; σ_v^2 is the variance of v_i; and τ is the vector representing the correlation between ε_i^* and v_i.

A Cholesky decomposition of the variance-covariance matrix of $(\varepsilon_i^{*\prime}, v_i)'$ gives:

$$
\begin{bmatrix} \varepsilon_i^* \\ v_i \end{bmatrix} = \begin{bmatrix} I_q & 0 \\ \sigma_v \tau' & \sigma_v \sqrt{1 - \tau' \tau} \end{bmatrix} \begin{bmatrix} \varepsilon_i^* \\ r_i^* \end{bmatrix},
\tag{2}
$$

where $r_i^* \sim \mathrm{N}(0,1)$, r_i^*, and ε_i^* are independent. Therefore, we have:

$$\begin{aligned} v_i &= \varepsilon_i^{*'}\sigma_v\rho + \sigma_r r_i^* \\ &= \varepsilon_i^{'}\eta + \sigma_r r_i^*, \end{aligned} \tag{3}$$

where $\sigma_r = \sigma_v\sqrt{1-\tau'\tau}$, $\eta = \sigma_r\Omega_\varepsilon^{-1/2}\tau \,/\, \sqrt{1-\tau'\tau}$, and $r_i = \sigma_r r_i^*$. Then, the frontier equation can be written as follows:

$$y_i = \rho\sum\nolimits_j w_{ij}y_j + x_{1i}^{'}\beta + (x_i - z_i\delta)'\eta + e_i, \tag{4}$$

where $e_i = r_i - u_i$ and $(x_i - z_i\delta)'\eta$ is a bias correction term. The density function of r_i is given by:

$$f_{r_i}(r_i) = \left(2\pi\sigma_r^2\right)^{-1/2}\exp\left(-\frac{r_i^2}{2\sigma_r^2}\right). \tag{5}$$

Moreover, the density function of ε_i is given by:

$$f_{\varepsilon_i}(\varepsilon_i) = \left|2\pi\Omega_\varepsilon\right|^{-1/2}\exp\left(-\frac{1}{2}tr(\Omega_\varepsilon^{-1}\varepsilon_i\varepsilon_i^{'})\right). \tag{6}$$

As ε_i and e_i are independent, the log-likelihood function is given by:

$$\ln L = \ln L_0 + \sum\nolimits_i \left(\ln L_{1i} + \ln L_{2i}\right), \tag{7}$$

where $\ln L_0 = \ln\left|I_N - \rho W\right|$ is the scaled logged determinant of the Jacobian of the transformation from e to y and W is the row-normalized matrix for weights with zero diagonals;[7]

[7] It is standard to assume in the literature that diagonal elements of W are zero. This rules out self influence possibility.

$$\ln L_{1i} = -\frac{1}{2}\ln(2\pi\sigma_r^2) - \frac{1}{2}\frac{e_i^2}{\sigma_r^2} + \frac{1}{2}\left(\frac{\mu_{i*}^2}{\sigma_{i*}^2} - \frac{\mu^2}{\sigma_u^2}\right) + \ln\left(\frac{\sigma_{i*}\Phi(\frac{\mu_{i*}}{\sigma_{i*}})}{\sigma_u\Phi(\frac{\mu}{\sigma_u})}\right);$$

$$\ln L_{2i} = -\frac{1}{2}\ \ln\ \left|2\pi\Omega_\varepsilon\right|\ + tr(\Omega_\varepsilon^{-1}\varepsilon_i\varepsilon_i')\ ;$$

$$\mu_i^* = \frac{-\sigma_u^2 e_i h_i + \mu\sigma_r^2}{\sigma_u^2 h_i^2 + \sigma_r^2}\ ;\ \ \sigma_{i*}^2 = \frac{\sigma_r^2\sigma_u^2}{\sigma_u^2 h_i^2 + \sigma_r^2}\ ;\ \ e_i = y_i - \rho\sum_j w_{ij}y_j - x_{1i}'\beta - \varepsilon_i'\eta\ ;\ \text{and}\ \ \varepsilon_i = x_i - z_i\delta\ .$$

By maximizing the total log-likelihood $\ln L$, we obtain the estimates for the model's parameters. Under standard conditions, the our estimator is consistent as $n \to \infty$.

One of the outstanding difficulties is that when we have a large sample, $\left|I_N - \rho W\right|$ term is the determinant of a large matrix, which needs to be re-calculated at each iteration of the optimization procedure. One potential possibility, as suggested by Pace an Perry (1997), is evaluating $\left|I_N - \rho W\right|$ term usig a vector of values for ρ in the internal $\left[\rho_{\min}, \rho_{\max}\right]$. These values need to be calculated before optimization and thus would only require calculation of the corresponding vector of determinants once. If we have a sufficiently fine grid of ρ values, we can use interpolated values of $\left|I_N - \rho W\right|$ to obtain intervening points.[8] In what follows, we assume that $\rho \in \left[0,1\right.$, the elements of W are non-negative, and all the diagonal elements of W are zero. An implication of this this assumption is that $\left|I_N - \rho W\right| \neq 0$ and thus $I_N - \rho W$ is non-singular. As mentioned by LeSage and Pace (1999), $\rho \in \left[0,1\right.$ assumption is widely employed in the literature.[9] Moreover, as described by Kutlu (2018) and we will argue later in the paper, this assumption is useful when interpreting the

[8] There are a number of approaches to obtain this determinant (computationally) efficiently. See LeSage and Pace (1999) for details of these approaches as well as numerical approaches used in the maximum likelihood estimation.

[9] Glass, Kenjegalieva, and Sickles (2014) assume that $\rho \in\ 1/\lambda_{\min}, 1$ where $\lambda_{\min}$ is the smallest real characteristic root of W.

efficiency estimates. Following Glass, Kenjegalieva, and Sickles (2014), we also assume that the rows and columns of W and $I_N - \rho W^{-1}$ are uniformly bounded in absolute value before row-normalizing W. This assumption implies that the spatial process for the dependent variable has a fading memory (Kelejian and Pruchas, 1998, 1999). The computational burden can be reduced further by applying variations of concentrated log-likelihood approaches in the literature (e.g., Elhorst, 2009; Glass, Kenjegalieva, and Sickles, 2014). Finally, note that when we have panel data, we can simply replace W by $I_T \otimes W$ where T is the number of time periods; and the rest of the analysis remains the same.

Once we obtain the parameter estimates, the inefficiency term u_i can be predicted via:

$$\hat{u}_i = E\left[u_i \mid e_i\right] = h_i \left[\mu_{i*} + \frac{\sigma_{i*}\phi(\frac{\mu_{i*}}{\tilde{\sigma}_{i*}})}{\Phi(\frac{\mu_{i*}}{\sigma_{i*}})}\right] \tag{8}$$

In practice, this equation is evaluated at $\hat{e}_i = y_i - \hat{\rho}\sum_j w_{ij} y_j - x_{1i}^{'}\hat{\beta}$.

2.2. Direct, Indirect, and Total Efficiency Estimates

As argued by LeSage (2009) the marginal effect of explanatory variables would be a function of the SAR term; and thereore the β parameter estimates cannot be interpreded as marginal effects. In matrix notation, the production equation is given by:

$$y = \rho W y + X_1 \beta - u + v\,, \tag{9}$$

which can be written as follows:

$$y = I_N - \rho W^{-1} X_1 \beta - I_N - \rho W^{-1} u + I_N - \rho W^{-1} v. \tag{10}$$

After renaming the variables, we have:

$$y = \tilde{X}_1 \beta - \tilde{u} + \tilde{v}, \tag{11}$$

where $\tilde{X}_1 = (I_N - \rho W)^{-1} X_1$, $\tilde{u} = (I_N - \rho W)^{-1} u$, and $\tilde{v} = (I_N - \rho W)^{-1} v$. Therefore, the marginal effects for Cobb-Douglas production function are given by:

$$\frac{\partial y_i}{\partial x_{1kj}} = \beta_k \left[(I_N - \rho W)^{-1} \right]_{ij}, \tag{12}$$

where x_{1kj} is the k^{th} frontier variable for productive unit j; β_k is the k^{th} component of β; and $\left[(I_N - \rho W)^{-1} \right]_{ij}$ is the ij^{th} element of $(I_N - \rho W)^{-1}$. The total marginal effect of x_{1k}, k^{th} frontier variable, is defined as the marginal change in y_i as a response to changes in x_{1kj} for all j:

$$\sum_j \frac{\partial y_i}{\partial x_{1kj}} = \beta_k \sum_j \left[(I_N - \rho W)^{-1} \right]_{ij}. \tag{13}$$

As pointed out by Kutlu (2018), the total inefficiency is captured by the $\tilde{u}$ term, not by u. Kutlu (2018) showed that when W is a row-normalized weighting matrix with diagonal elements being zero and $\rho \in [0,1)$, we have $(I_N - \rho W)^{-1} \geq 0$, i.e., all elements are non-negative. Therefore, $\tilde{u}_i \geq 0$, $\tilde{u}_i$ is a non-decreasing function of components of u, and if $u = 0$, then $\tilde{u}_i = 0$. These imply, as argued by Kutlu (2018), that we can use $\tilde{u}_i = 0$ to represent the full efficiency benchmark.

In our spatial model, the usual formula for calculating (total) efficiency is not valid as it ignores the spatial spillovers. The corrected efficiency can be calculated by:

$$E_i = exp(-\tilde{u}_i). \tag{14}$$

This is a generalization of the usual formula as when $\rho = 0$, we have $\tilde{u}_i = u_i$. The shares of direct and indirect inefficiencies (see Kutlu, 2018) are given by:

$$SIE_i^{dir} = \frac{\left[(I_N - \rho W)^{-1} u \right]_{ii}}{\tilde{u}_i}$$
$$SIE_i^{ind} = \frac{\sum_{i \neq j} \left[(I_N - \rho W)^{-1} u \right]_{ij}}{\tilde{u}_i}. \tag{15}$$

These shares can be used to decompose efficiency into direct and indirect efficiency components.

2.3. Testing Endogeneity

Amsler, Prokhorov, and Schmidt (2016) and Karakaplan and Kutlu (2017a,b) describe a simple test, using similar ideas with the Durbin-Wu-Hausman test, for endogeneity for the non-spatial stochastic frontier models. These tests are applicable in our setting as well. We can test the endogeneity using the F-test for $\eta = 0$. If all components of η are jointly significant, we conclude that the bias correction term is needed and thus we have endogeneity. We can also check test the endogeneity of individual variables by testing the significance of the corresponding component of η.

3. Monte Carlo Simulations

To evaluate the performance of our proposed estimator in finite samples, we conduct a small Monte Carlo experiment. We consider the following data generating process (DGP):

$$\begin{aligned} y_i &= \rho \sum_{j=1}^{n} w_{ij} y_j + z_{1i}\alpha + q_{fi}\beta + v_i - u_i \\ u_i &= h_i u_i^* \\ h_i &= \left[\exp(z_{2i}\varphi_1 + q_{ui}\varphi_2)\right]^{1/2} \\ u_i^* &\sim N^+(0, \exp(c_u)) \end{aligned} ,$$

where z_{1i} and z_{2i} are exogenous variables, v_i and u_i^* are independent random variables. The spatial weights w_{ij} are generated using row normalized exponential distance weights of the form:

$$w_{ij} = \begin{cases} \exp(-d_{ij}) \,/\, \sum_{i \neq j} \exp(-d_{ij}), & i \neq j \\ 0 \quad , & i = j \end{cases}$$

where d_{ij} are the *centroid distances* between each pair of spatial units i and j. The endogenous variables q_{fi} and q_{ui} are generated as follows:

$$q_{fi} = z_{3i}\delta_1 + \varepsilon_{1i}$$
$$q_{ui} = z_{4i}\delta_2 + \varepsilon_{2i}$$

where $z_i = (z_{1i}, z_{2i}, z_{3i}, z_{4i})' \sim MN(0,\Sigma)$, and the correlation among q_{fi}, q_{ui}, and v_i are genenerated via:

$$\begin{pmatrix} \varepsilon_i \\ v_i \end{pmatrix} \sim \mathbf{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon_1}^2 & 0 & \sigma_{\varepsilon_1}\sigma_v\tau_1 \\ 0 & \sigma_{\varepsilon_2}^2 & \sigma_{\varepsilon_2}\sigma_v\tau_2 \\ \sigma_{\varepsilon_1}\sigma_v\tau_1 & \sigma_{\varepsilon_2}\sigma_v\tau_2 & \sigma_v^2 \end{pmatrix}\right)$$

where $\varepsilon_i = (\varepsilon_{1i}, \varepsilon_{2i})'$. Note that when $\tau_1 = 0$, q_{fi} becomes exogenous and likewise, when $\tau_2 = 0$, q_{ui} becomes exogenous. Finally, when $\tau_1 = \tau_2 = 0$, all variables are exogenous and this is our branchmark or "exogenous model." In each experiment, we consider the following values for $(\tau_1, \tau_2) = \{(0,0),(0.7,0.7)\}$, and we fixed the values of $\rho = 0.7$, $\alpha = \beta = 0.75$, $\delta_1 = \delta_2 = 1$, $c_u = -0.5$, $\sigma_v = \sqrt{0.1}$, $\sigma_{\varepsilon_1} = 0.5$, $\sigma_{\varepsilon_2} = 0.75$, $\varphi_1 = 0.4$, $\varphi_2 = 0.9$, and

$$\Sigma = \begin{bmatrix} 1 & 0.2 & 0.2 & 0.2 \\ 0.2 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.2 \\ 0.2 & 0.2 & 0.2 & 1 \end{bmatrix}.$$

Finally, we consider the following sample sizes: $n = (100, 200, 400)$, and the Monte Carlo experiments are conducted with 500 replications. We announce the biases and root mean squared errors for the parameter estimates and efficiency estimates as well as spearman correlation of true and estimated efficiencies.

4. Empirical Example

4.1. Data

4.2. Empirical Model and Estimation Results

5. Conclusion

The conventional stochastic frontier models neither allow spatial spillovers nor endogeneity. If any of the frontier or environmental variables are correlated with the two-sided error term; or the SAR component is omitted while being a relevant term, then parameter and efficiency estimates would be inconsistent. We presented the first model that can solve both issues simultaneously by employing a control function approach.

XXXXX

References

Aigner, D.J., Lovell, C.A.K., and Schmidt P. (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics*, 6, 21-37.

Almanidis, P. (2013), Accounting for Heterogeneous Technologies in the Banking Industry: A Time-Varying Stochastic Frontier Model with Threshold Effects, *Journal of Productivity Analysis*, 39, 191-205.

Almanidis, P., Karakaplan, M.U., and Kutlu, L. (2016), Are We Looking at the Wrong End of the Bank Size Spectrum, Working Paper.

Almanidis P, Qian J, and Sickles R. (2014), Stochastic Frontiers with Bounded Inefficiency. Festschrift in honor of Peter Schmidt: Econometric Methods and Applications. Berlin: Springer Science and Business Media.

Amsler, C., Prokhorov, A., and Schmidt, P. (2016), Endogenous Stochastic Frontier Models, *Journal of Econometrics*, 190, 280-288.

Amsler, C., Prokhorov, A., and Schmidt, P. (2017), Endogenous Environmental Variables in Stochastic

Frontier Models, *Journal of Econometrics*, 199, 131-140.

Battese, G.E. and Coelli, T.J. (1992), Frontier Production Functions, Technical Efficiency and Panel Data with Application to Paddy Farmers in India, *Journal of Productivity Analysis*, 3, 153-169.

Battese, G.E. and Corra, G.S. (1977), Estimation of a Production Frontier Model: With a Generalized Frontier Production Function and Panel Data, *Australian Journal of Agricultural Economics*, 21, 169-179.

Belotti, F. and Ilardi, G. (2018), Consistent Inference in Fixed-Effects Stochastic Frontier Models, *Journal of Econometrics*, 202, 161-177.

Berger, A.N. (1993), Distribution Free Estimates of Efficiency in US Banking Industry and Tests of the Standard Distributional Assumptions, *Journal of Productivity Analysis*, 4, 261-292.

Berger, A.N. and Hannan, T.H. (1998), The Efficiency Cost of Market Power in the Banking Industry: A Test of the "Quiet Life" and Related Hypotheses, *Review of Economics and Statistics*, 80, 454-465.

Bresnahan, T.F. (1989), Empirical Studies of Industries with Market Power, *Handbook of industrial organization* 2: 1011-1057.

Caudill, S.B. and Ford, J. (1993), Biases in Frontier Estimation Due to Heteroscedasticity, *Economic Letters*, 41, 17-20.

Chen, Y-Y, Schmidt, P., and Wang H-J. (2014), Consistent Estimation of the Fixed Effects Stochastic Frontier Model, *Journal of Econometrics*, 181, 65-76.

Colombi, R., Kumbhakar, S., Martini, G., and Vittandini, G. (2014), Closed-Skew Normality in Stochastic Frontiers with Individual Effects and Long/Short-Run Efficiency, *Journal of Productivity Analysis*, 42, 123-136.

Cornwell C., Schmidt P., and Sickles R.C. (1990), Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels, *Journal of Econometrics*, 46, 185-200.

Demsetz, H. (1973), Industry Structure, Market Rivalry, and Public Policy, *Journal of Law and Economics*, 16, 1-9.

Greene, W.H. (1980a), Maximum Likelihood Estimation of Econometric Frontier Functions, *Journal of Econometrics*, 3, 27-56.

Greene, W.H. (1980b), On the Estimation of a Flexible Frontier Production Model, *Journal of Econometrics*, 13, 101-115.

Greene, W.H. (2003), Simulated Likelihood Estimation of the Normal-Gamma Stochastic Frontier Function, *Journal of Productivity Analysis*, 19, 179-190.

Greene, W.H. (2005a), Fixed and Random Effects in Stochastic Frontier Models, *Journal of Productivity Analysis*, 23, 7-32.

Greene, W.H. (2005b), Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model, *Journal of Econometrics*, 126, 269-303.

Greene, W.H. (2008), Econometric analysis, 6th edition. Englewood Cliffs, N. J.: Prentice Hall.

Griffiths, W.E. and Hajargasht, G. (2016), Some Models for Stochastic Frontiers with Endogeneity, *Journal of Econometrics*, 190, 341-348.

Griliches, Z. and Mairesse, J. (1998), Production Functions: The Search for Identification. In: Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium, ed. S. Strøm. Cambridge, UK: Cambridge University Press.

Guan Z., Kumbhakar S.C., Myers R.J., and Lansink A.O. (2009), Measuring Excess Capital Capacity in Agricultural Production, *American Journal of Agricultural Economics*, 91, 765-776.

Hicks, J.R. (1935), Annual Survey of Economic Theory: The Theory of Monopoly, *Econometrica*, 3, 1-20.

Jamasb, T., Orea, L. and Pollitt, M. (2012), Estimating Marginal Cost of Quality Improvements: The Case of the U.K. Electricity Distribution Companies, *Energy Economics*, 34, 1498-1506.

Jayaratne, J. and Strahan, P.E. (1996), The Finance-Growth Nexus: Evidence from Bank Branch Deregulation, *Quarterly Journal of Economics*, 111, 639–670.

Karakaplan, M.U., and Kutlu, L. (2017a), Handling Endogeneity in Stochastic Frontier Analysis, *Economics Bulletin,* 37, 889-901.

Karakaplan, M.U. and Kutlu, L. (2017b), Endogeneity in Panel Stochastic Frontier Models: An Application to the Japanese Cotton Spinning Industry, *Applied Economics*, 49, 5935-5939.

Karakaplan, M.U. and Kutlu, L. (2018a), Estimating Market Power Using a Composed Error Model, Forthcoming in *Scottish Journal of Political Economy*.

Karakaplan, M.U. and Kutlu, L. (2018b), School District Consolidation Policies: Endogenous Cost Inefficiency and Saving Reversals, Forthcoming in *Empirical Economics*.

Karmalkar, A.V. and Bradley, R.S. (2017), Consequences of Global Warming of 1.5 ˚C and 2 ˚C for Regional Temperature and Precipitation Changes in the Contiguous United States, *PLoS ONE*, 12(1): e0168697. doi:10.1371/journal. pone.0168697.

Khatri, C.G. (1968), Some Results for the Singular Normal Multivariate Regression Models., *Sankhya*, 30, 267-280.

Koetter, M., Kolari, J.W., and Spierdijk, L. (2012) Enjoying the Quiet Life under Deregulation? Evidence from Adjusted Lerner Indices for U.S. Banks, *Review of Economics and Statistics*, 94, 462-480.

Kroszner, R.S. and Strahan, P.E. (1999), What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions, *Quarterly Journal of Economics*, 104, 1437-1467.

Kumbhakar, S.C. (1990), Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency, *Journal of Econometrics*, 46, 201-211.

Kumbhakar, S., Lien, G., and Hardaker, J.B., (2014), Technical Efficiency in Competing Panel Data Models: A Study of Norwegian Grain Farming, *Journal of Productivity Analysis*, 41, 321-337.

Kutlu, L. (2010), Battese-Coelli Estimator with Endogenous Regressors, *Economics Letters*, 109, 79-81.

Kutlu, L. (2012), U.S. Banking Efficiency, 1984-1995, *Economics Letters*, 117, 53-56.

Kutlu, L. (2017), A Constrained State Space Approach for Estimating Firm Efficiency, *Economics Letters*, 152, 54-56.

Kutlu, L. (2018), A Distribution-Free Stochastic Frontier Model with Endogenous Regressors, *Economics Letters*, 152, 54-56.

Kutlu, L. and Nair-Reichert, U. (2018a), Agglomeration Effects and Spatial Spillovers in Efficiency Analysis – A Distribution-Free Methodology, *Unpublished manuscript*.

Kutlu, L. and Nair-Reichert, U. (2018b), Efficiency and Executive Pay: Evidence from Indian Manufacturing Sector, *Unpublished manuscript*.

Kutlu, L. and Sickles, C.R. (2012), Estimation of Market Power in the Presence of Firm Level Inefficiencies, *Journal of Econometrics*, 168, 141-155.

Kutlu, L. and Wang, R. (2018), Estimation of Cost Efficiency without Cost Data, *Journal of Productivity Analysis*, 49, 137-151.

Meeusen, W. and Van Den Broeck J. (1977), Efficiency Estimation From Cobb-Douglas Production Function with Composed Errors, *International Economic Review*, 18, 435-444.

Mutter, R.L., Greene, W.H., Spector, W., Rosko, M.D., and Mukamel, D.B. (2013), Investigating the Impact of Endogeneity on Inefficiency Estimates in the Application of Stochastic Frontier Analysis to Nursing Homes, *Journal of Productivity Analysis*, 39, 101-110.

Murphy, K.M. and Topel R.H. (1985), Estimation and Inference in Two-Step Econometric Models, *Journal of Business and Economic Statistics*, 3, 370-379.

Nelsen, R.B. (2006), An Introduction to Copulas, Vol. 139 of Springer Series in Statistics, 2nd ed. New York: Springer.

Orea, L. and Steinbuks, J. (2018), Estimating Market Power in Homogenous Product Markets Using a Composed Error Model: Application to the California Electricity Market, *Economic Inquiry*, 56, 1296-1321.

Park, S. and Gupta, S. (2012), Handling Endogenous Regressors by Joint Estimation Using Copulas, *Marketing Science*, 31, 567-586.

Perlof, J.M., Karp, L.S., and Golan, A. (2007). Estimating Market Power and Strategies. Cambridge University Press.

Reifschneider, D. and Stevenson, R. (1991), Systematic Departures From the Frontier: A Framework for the Analysis of Firm Inefficiency, *International Economic Review*, 32, 715-723.

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

Shee, A. and Stefanou, S.E. (2014), Endogeneity Corrected Stochastic Production Frontier and Technical Efficiency, *American Journal of Agricultural Economics*, aau083.

Stevenson, R. (1980), Likelihood Functions for Generalized Stochastic Frontier Functions, *Journal of Econometrics*, 13, 57-66.

Tran, K.C. and Tsionas, E.G. (2015), Endogeneity in Stochastic Frontier Models: Copula Approach without External Instruments, *Economics Letters*, 133, 85-88.

Tran, K.C. and Tsionas, E.G. (2013), GMM Estimation of Stochastic Frontier Model with Endogenous Regressors, *Economics Letters,* 118, 233-236.

Tsionas, E.G. (2002), Stochastic Frontier Models with Random Coefficients, *Journal of Applied Econometrics*, 17, 121-47.

Wang, H.J. and Ho, C.W. (2010), Estimating Fixed-Effect Panel Stochastic Frontier Models by Model Transformation, *Journal of Econometrics*, 157, 286-296.

Wang, H.J. and Schmidt, P. (2002), One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels, *Journal of Productivity Analysis*, 18, 129-144.