2-2018

# Estimating efficiency in a spatial autoregressive stochastic frontier model

Levent Kutlu
*The University of Texas Rio Grande Valley*, levent.kutlu@utrgv.edu

# Estimating Efficiency in a Spatial Autoregressive Stochastic Frontier Model

Levent Kutlu[a]

## Abstract

The spatial autoregressive stochastic frontier model of Glass, Kenjegalieva, and Sickles (2016) is based on distributional assumptions on two-sided and one-sided error terms. After estimating the model parameters, the efficiency estimates need to be corrected due to the presence of spatial autoregressive term in their model. Glass, Kenjegalieva, and Sickles (2016) estimate the corrected efficiencies by employing ideas from a distribution-free method on the efficiency estimation, which may be sensitive to outliers. We propose an alternative way to correct efficiency estimates that is in line with the distribution-based methods.

**Keywords:** Efficiency; Spatial autoregression; Stochastic frontier

**JEL Classification:** C23; D24

[a] School of Economics, Georgia Institute of Technology, Atlanta, GA, USA.
Email: levent.kutlu@econ.gatech.edu, phone: +1-404-385-1363

## 1. Introduction

In cross-sectional or panel data stochastic frontier models, the omission of spatial lagged dependent variable, which captures spatial autoregressive (SAR) dependence, may lead to bias in parameter and efficiency estimates if spatial dependence is present.[1] The literature on such spatial stochastic frontier models is sparse. Glass, Kenjegalieva, and Paez-Farrell (2013) and Glass, Kenjegalieva, and Sickles (2014) exemplify recently developed distribution-free SAR stochastic frontier models where efficiencies are calculated by using the method of Schmidt and Sickles (1984).[2] In contrast to these distribution-free studies, Glass, Kenjegalieva, and Sickles (2016) (GKS) propose a SAR stochastic frontier model where efficiency is calculated from a composed error structure by making distributional assumptions on the two-sided and one-sided error terms.[3] GKS argue that the coefficients in the frontier cannot be interpreted as elasticities because the marginal effect of an independent variable in the frontier is a function of the SAR variable. They solve this problem by disentangling the effect of independent variable from the effect of SAR variable. The estimation of (total) efficiency involves a similar problem.[4] However, in order to estimate the SAR-corrected efficiencies, although they assume the half-normal distribution for the one-sided error term, they adapt the method of Schmidt and Sickles (1984). Therefore, their efficiency estimates have the combined weaknesses of distribution-based and distribution-free models. For example, the efficiency estimates obtained by distribution-free approach may be sensitive to outliers. For this reason, in practice some researchers trim the estimated individual

[1] For non-frontier literature on spatial models, see Baltagi (2011, 2013).
[2] See also Cornwell, Schmidt, and Sickles (1990).
[3] See Tsionas and Michaelides (2016) for another study that relies on distributional assumptions where inefficiency term is spatial autoregressive in the Bayesian econometrics context.
[4] Total efficiency includes spillover effects for efficiency. We will define total efficiency later in the paper.

effects terms from bottom and top to avoid outliers.[5] We propose an alternative way to obtain SAR-corrected efficiency estimates that is in line with the distribution-based methods, which is not subject to this criticism. Moreover, we decompose the total inefficiency into firm-specific inefficiency spillover components.

In the next section, we describe the SAR stochastic frontier model and our estimation procedure for efficiency. Then, we make our concluding remarks.

## 2. Efficiency Estimation in a SAR Stochastic Frontier Model

For the sake of fixing ideas, we consider production function estimation. The model of GKS is given by: [6]

$$y_t = \rho W y_t + X_t \beta - u_t + v_t \tag{1}$$

where $y_t = \left(y_{1t}, y_{2t}, ..., y_{Nt}\right)'$ is the $T \times 1$ vector of the logarithm of output at time $t$ where $T$ is the number of time periods and $N$ is the number of firms; $\rho$ is a spatial correlation parameter; $W$ is the $N \times N$ spatial row-normalized weighting matrix; $X_t$ is the $N \times K$ matrix of frontier variables at time $t$; $u_t = \left(u_{1t}, u_{2t}, ..., u_{Nt}\right)'$ is the $N \times 1$ vector of non-negative inefficiency term at time $t$ where $u_{it}$ has half-normal distribution; and $v_t = \left(v_{1t}, v_{2t}, ..., v_{Nt}\right)'$ is the $N \times 1$ vector of usual two-sided error term with independent components at time $t$, which is independent from $u_t$. They rewrite the model as:

$$\left(I_N - \rho W\right) y_t = X_t \beta - u_t + v_t \tag{2}$$

[5] See Kutlu (2012) and Duygun, Kutlu, and Sickles (2016) for studies that obtain efficiency estimates by trimming. Also, see Kutlu (2017) for a solution to outlier issues in distribution-free models.
[6] See Adetutu et al. (2015) for a stochastic frontier model of spatial interaction where the spatial dependence is captured by exogenous explanatory variables.

and transform it by $S(\rho) = (I_N - \rho W)^{-1}$ so that:

$$y_t = S(\rho) X_t \beta - S(\rho) u_t + S(\rho) v_t. \tag{3}$$

In order to estimate the parameters of this model, they first obtain a conditional log-likelihood that has $\rho$ as the only unknown parameter.[7] They recover $\beta$ conditional on the estimate of $\rho$. Then, they estimate the $u_{it}$ term by a pseudo-likelihood estimation method.[8]

The estimates for $\beta$ cannot be interpreted as marginal effects. GKS calculate the marginal effects as follows:

$$\frac{\partial y_{it}}{\partial x_{jt,k}} = \beta_k \left[ (I_N - \rho W)^{-1} \right]_{ij} \tag{4}$$

where $x_{jt,k}$ is the $k^{th}$ explanatory variable for firm $j$ at time $t$; $\beta_k$ is the $k^{th}$ component of $\beta$; and $\left[ (I_N - \rho W)^{-1} \right]_{ij}$ is the $ij^{th}$ element of $(I_N - \rho W)^{-1}$.[9] The total marginal effect of $k^{th}$ explanatory variable at time $t$ is defined as the marginal change in $y_{it}$ in response to changes in $x_{jt,k}$'s for all firms at time $t$. That is, the total marginal effect of $k^{th}$ explanatory variable is given by:

$$\sum_j \frac{\partial y_{it}}{\partial x_{jt,k}} = \sum_j \beta_k \left[ (I_N - \rho W)^{-1} \right]_{ij}. \tag{5}$$

[7] They follow Elhorst (2009) when deriving the conditional log-likelihood.
[8] See Han, Ryu, and Sickles (2016) for a similar methodology applied to the case where the weighting matrix is time-varying.
[9] GKS suggest using mean total marginal effect in respond to changing an explanatory variable.

In a SAR stochastic frontier model, the standard formula for calculating (total) efficiency, i.e., $\xi_{it} = \exp(-u_{it})$, is not correct and must be corrected in a way similar to total marginal effects for explanatory variables. In particular, as suggested by GKS, the total $u_t$ (total inefficiency) can be calculated by $u_t^{tot} = (I_N - \rho W)^{-1} u_t$. Alternatively, GKS consider the total $\xi_t$ (total efficiency) by $\xi_t^{tot} = (I_N - \rho W)^{-1} \xi_t$. They argue, however, that while $\xi_{it}$ is relative to an absolute best practice frontier, $\xi_{it}^{tot}$ may not be relative to the absolute best frontier. Similarly, for an arbitrary *W* matrix, it is not clear whether $u_{it}^{tot}$ or $\xi_{it}^{tot}$ would be non-negative, which is a necessary condition for total inefficiency and total efficiency being well-defined. In order to address this concern, they adapt the Schmidt and Sickles (1984) method and apply it to $\xi_t$.[10] In particular, they calculate the SAR-corrected total efficiencies at time $t$ as follows:[11]

$$\tilde{E}_{it}^{tot} = \frac{\sum_j \left[ (I_N - \rho W)^{-1} \xi_t \right]_{ij}}{\max_i \left( \sum_j \left[ (I_N - \rho W)^{-1} \xi_t \right]_{ij} \right)} \tag{6}$$

which measures the efficiency spillovers to the $i^{th}$ unit from all the $j^{th}$ units. As we stated earlier, this measure may be sensitive to outliers. Moreover, since $\xi_t$ is estimated by a distribution-based method, the total efficiency estimates based on Equation (6) are subject to criticisms for both distribution-based and distribution-free approaches.

We propose estimating the total efficiency of firm $i$ at time $t$ as follows:

$$E_{it}^{tot} = \exp\left(-u_{it}^{tot}\right) \tag{7}$$

[10] GKS also consider direct and indirect efficiencies, which constitute a decomposition of the total inefficiency. They also consider asymmetric efficiency spillovers. Our method can be adapted to these efficiency concepts as well.
[11] In the terminology of GKS, SAR-corrected total efficiency is relative total efficiency.

which can be considered as a generalization of distribution-based efficiency estimates for SAR stochastic frontier models. The following theorem illustrates why the total efficiency formula in Equation (7) is reasonable.

**Theorem:** Let $W$ be a row-normalized weighting matrix with diagonal elements being equal to zero and $0 \le \rho < 1$. Then, all elements of $\left(I_N - \rho W\right)^{-1}$ are non-negative, i.e., $\left[\left(I_N - \rho W\right)^{-1} \xi_t\right]_{ij} \ge 0$ for all $i$ and $j$.

**Proof:** *We know that the inverse of an M-matrix has non-negative elements where an M-matrix is a matrix with non-positive off-diagonal elements with eigenvalues whose real parts are positive. First, note that the non-diagonal elements of $I_N - \rho W$ are non-positive. Since the absolute value of eigenvalues of a row-normalized matrix is smaller or equal to 1 and the eigenvalues of $I_N - \rho W$ are $\{1-\rho\lambda_1, 1-\rho\lambda_2, ..., 1-\rho\lambda_N\}$ where $\lambda_i$'s are eigenvalues of $W$, the real part of eigenvalues of $I_N - \rho W$ are positive. Therefore, $I_N - \rho W$ is an M-matrix, which completes the proof.*

An immediate implication of this theorem is that whenever $0 \le \rho < 1$, which is a condition that we expect to hold in most occasions in practice, we have $u_{it}^{tot} \ge 0$. Moreover, $u_{it}^{tot}$ is a non-decreasing function of components of $u_t$ and $u_{it}^{tot} = 0$ if $u_t = 0$. Hence, as in the conventional distribution-based stochastic frontier models, we can use $u_{it}^{tot} = 0$ as a benchmark to represent the full efficiency. Moreover, note that Equation (3) can be re-written as:

$$y_t = \tilde{X}_t \beta - \tilde{u}_t + \tilde{v}_t. \tag{8}$$

where $\tilde{X}_t = S(\rho) X_t$, $\tilde{u}_t = S(\rho) u_t = u_t^{tot}$, and $\tilde{v}_t = S(\rho) v_t$. Therefore, Equation (7), $E_{it} = \exp\left(-u_{it}^{tot}\right) = \exp\left(-\tilde{u}_{it}\right)$, is internally consistent with Equation (3) and can be used as a generalization of the efficiency formula for conventional stochastic frontier models without the SAR term, i.e., $\exp\left(-u_{it}^{tot}\right) = \exp\left(-u_{it}\right)$ if $\rho = 0$. A difficulty in SAR models (frontier or non-frontier) is that when $\rho \to 0$, the mean marginal effects converge to zero. Similarly our total efficiency estimates based on Equation (7) would converge to zero as $\rho \to 0$.

Observe that the efficiency calculation based on $u_t^{tot} = \left(I_N - \rho W\right)^{-1} u_t$ would have $\left(I_N - \rho W\right)^{-1}$ inside the exponential operator, i.e., $\exp\left(-u_t^{tot}\right) = \exp\left(-\left(I_N - \rho W\right)^{-1} u_t\right)$. Whereas the efficiency calculation based on $\xi_t^{tot} = \left(I_N - \rho W\right)^{-1} \xi_t$ would have $\left(I_N - \rho W\right)^{-1}$ outside the exponential operator, i.e., $\xi_t^{tot} = \left(I_N - \rho W\right)^{-1} \xi_t = \left(I_N - \rho W\right)^{-1} \exp\left(-u_t\right)$. A potential distinction between these calculation methods is that compared to the latter method the former method may lead the median and minimum total efficiencies to be pulled away from the highest total efficiency.[12]

We can calculate the firm-specific efficiency spillover of firm $j$ on firm $i$ as follows:

$$E_{it}^{j} = \exp\left(-\left[\left(I_N - \rho W\right)^{-1} u_t\right]_{ij}\right). \tag{9}$$

Using similar ideas, it is possible to obtain a decomposition of total inefficiency so that the share of the inefficiency spillover of firm $j$ on firm $i$ at time $t$ is given by:

$$SIE_{it}^{j} = \frac{\left[\left(I_N - \rho W\right)^{-1} u_t\right]_{ij}}{u_{it}^{tot}}. \tag{10}$$

[12] We thank the anonymous referee for pointing this out.

Therefore, the shares of direct and indirect inefficiencies are given by:

$$SIE_{it}^{dir} = \frac{\left[ \left( I_N - \rho W \right)^{-1} u_t \right]_{ii}}{u_{it}^{tot}}$$
$$SIE_{it}^{ind} = \frac{\sum_{i \neq j} \left[ \left( I_N - \rho W \right)^{-1} u_t \right]_{ij}}{u_{it}^{tot}} \tag{11}$$

In contrast to the decomposition of total efficiency suggested by GSK, our decomposition of inefficiency given in Equation (11) is exact so that so that $SIE_{it}^{dir} + SIE_{it}^{ind} = 1$. Moreover, our decomposition based on Equation (10) is finer in the sense that it is firm-specific. However, inefficiency concept has negative connotations and efficiency measures are more intuitive and therefore easier to interpret. Hence, both efficiency and inefficiency based decompositions have merits and may be considered as alternatives.

## 4. Conclusion

We provided an alternative way to estimate total efficiency for the SAR stochastic frontier model proposed by GKS. We also proposed a way to estimate firm-specific efficiency spillovers, which is useful in understanding the firm-specific spatial connections. To our knowledge, our paper is the first study that provides an exact decomposition of total inefficiency with firm-specific inefficiency spillover components. An advantage of our estimation procedure is that it is considerably easier to conduct the inefficiency analysis than it is to conduct the analysis in terms of efficiency as suggested by GKS. Moreover, our estimation procedure is consistent with the distribution-based methodology that is used by GKS and is more robust to outlier related issues when calculating efficiencies. Overall, both methods have merits and are alternative ways to calculate efficiency.

**References**


Adetutu, M., Glass, A.J., Kenjegalieva, K., and Sickles, R.C. (2015), The Effects of Efficiency and TFP Growth on Pollution in Europe: A Multistage Spatial Analysis, *Journal of Productivity Analysis*, 43, 307-326.

Baltagi, B.H. (2011), Spatial panels. In: Ullah, A., Giles, D.E.A. (Eds.), Handbook of Empirical Economics and Finance. Chapman and Hall, Taylor and Francis Group, Boca Raton, Florida.

Baltagi, B.H. (2013), Econometric Analysis of Panel Data, fifth ed. Wiley, Chichester, UK.

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

Duygun, M., Kutlu, L., and Sickles, R.C. (2016), Measuring Productivity and Efficiency: A Kalman Filter Approach, *Journal of Productivity Analysis*, 46, 155-167.

Elhorst, J.P. (2009), Spatial Panel Data Models. In: Fischer, M.M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis. Springer, New York.

Glass, A.J., Kenjegalieva, K., and Paez-Farrell, J. (2013), Productivity Growth Decomposition Using a Spatial Autoregressive Frontier Model, *Economics Letters*, 119, 291-295.

Glass, A.J., Kenjegalieva, K., and R.C. Sickles (2014), Estimating Efficiency Spillovers with State Level Evidence for Manufacturing in the US, *Economics Letters*, 123, 154-159.

Glass, A.J., Kenjegalieva, K., and R.C. Sickles (2016), A Spatial Autoregressive Stochastic Frontier Model for Panel Data with Asymmetric Efficiency Spillovers, *Journal of Econometrics* 190, 289-300.

Han, J., Ryu, D., and Sickles, R.C. (2016), Spillover Effects of Public Capital Stock Using Spatial Frontier Analyses: A First Look at the Data. In: Greene, W.H., Khalaf, L., Sickles, R.C., Veall, M., and M.C. Voia (Eds.), Productivity and Efficiency Analysis. Springer, Switzerland.

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

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

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

Tsionas, E.G. and Michaelides, P.G. (2016), A Spatial Stochastic Frontier Model with Spillovers: Evidence for Italian Regions, *Scottish Journal of Political Economy*, 63, 243-257.