The Spatial Efficiency Multiplier and Random Effects in Spatial Stochastic Frontier Models

Anthony J. Glass*, Karligash Kenjegalieva†, Robin C. Sickles‡§ and Thomas Weyman-Jones¶

November 2016

Abstract

We extend the emerging literature on spatial frontier models in three respects. Firstly, we account for latent heterogeneity by developing a maximum likelihood random effects spatial autoregressive (SAR) stochastic frontier model. Secondly, to analyze the finite sample properties of a spatial stochastic frontier model we develop a Monte Carlo experimental methodology which we then apply. Thirdly, we introduce the concept of the spatial efficiency multiplier and show that the efficiency benchmark for a productive unit from the structural form of a spatial stochastic frontier model differs from the efficiency benchmark from the reduced form of the model.

Key words: Stochastic frontier analysis (SFA); Spatial autoregression (SAR); Panel data; Random effects; Efficiency spillovers; Agriculture.

JEL Classification: C23; C51; D24; Q10.

*School of Business and Economics and Centre for Productivity and Performance, Loughborough University, Leics, UK, LE11 3TU. Email: A.J.Glass@lboro.ac.uk.

†School of Business and Economics and Centre for Productivity and Performance, Loughborough University, Leics, UK, LE11 3TU. Email: K.A.Kenjegalieva@lboro.ac.uk.

‡Corresponding author

§Department of Economics, Rice University, Houston, U.S., and School of Business and Economics and Centre for Productivity and Performance, Loughborough University, Leics, UK, LE11 3TU. Email: rsickles@rice.edu

¶School of Business and Economics and Centre for Productivity and Performance, Loughborough University, Leics, UK, LE11 3TU. Email: T.G.Weyman-Jones@lboro.ac.uk.

1 Introduction

Omitted variable bias in cross-sectional and panel data modeling due to a misspecification of spillovers from a spatial lag of the dependent variable, which captures what is referred to as spatial autoregressive (SAR) cross-sectional dependence, is well-established. This omitted variable bias was a key motivation for the development of the SAR model in seminal work by Cliff and Ord (1973; 1981). In a stochastic frontier framework, which is characterized by a composed disturbance with error and inefficiency components, efficiency measurement is correspondingly impacted by such misspecification resulting in biased efficiency score estimates. In a recent study Glass *et al.* (2016) (GKS from hereon) develop a maximum likelihood (ML) estimator of the SAR stochastic frontier model to account for this bias. To distinguish between the idiosyncratic error and inefficiency components of the composed disturbance GKS make distributional assumptions. Since spatial stochastic frontier modeling is an emerging literature we extend the basic GKS analysis in a number of ways.

Our first extension is to account for latent heterogeneity, which was not pursued in the GKS model, by developing a ML estimator of the random effects SAR stochastic frontier model. Unlike GKS who only estimate time-varying efficiency, which we label net time-variant efficiency (NVE) to highlight that it is net of invariance over time, we also estimate net time-invariant efficiency (NIE). We refer to the combined efficiency as gross time-variant efficiency ($GVE = NVE * NIE$). To also estimate NIE we use a three-step estimation procedure. In the first step we estimate the one-way SAR random effects model which distinguishes between the time-invariant and time-variant components of the composed error. The second step splits the time-variant component into the idiosyncratic error and net time-variant inefficiency (NVI) from which we obtain NVE. Similarly, the third step splits the time-invariant component into a time-invariant random error, which captures latent heterogeneity, and net time-invariant inefficiency (NII), which we transform to obtain NIE.[1]

More generally, there is a small body of literature on spatial frontier modeling which adopts a different approach to the one we utilize. In this literature one-way spatial panel models are estimated and efficiency is computed using the cross-sectional specific effects. Druska and Horrace (2004) is the first such study. By extending the cross-sectional spatial error model in Kelejian and Prucha (1999), they develop a GMM stochastic frontier model with fixed effects. Using the fixed effects they calculate time-invariant efficiency by applying the Schmidt and Sickles (1984) (SS from hereon) efficiency estimator, which

[1] NVE and GVE here should not be confused with net and gross efficiencies in Coelli *et al.* (1999). In their model net inefficiency omits the effects of exogenous firm specific variables (e.g., a policy variable) on mean inefficiency and gross inefficiency incorporates these effects. Allowing time-invariant variables to affect mean NII or time-variant variables to affect mean NVI is not pursued in this paper but is a logical extension of our SAR stochastic frontier model.

assumes a composed disturbance structure with idiosyncratic error and time-invariant inefficiency components. Glass *et al.* (2013) extended this literature by using the cross-sectional specific effects from a one-way spatial panel model to calculate time-variant efficiency using the time-variant extension of the SS estimator in Cornwell *et al.* (1990).

In contrast to the spatial stochastic frontier models that compute efficiency using the cross-sectional specific effects we follow GKS and compute efficiency by making distributional assumptions to distinguish between the idiosyncratic error and inefficiency components of the composed disturbance. Making such distributional assumptions is consistent with much of the non-spatial stochastic frontier literature including the canonical model of Aigner, Lovell and Schmidt (1977) (ALS from hereon). Our model is therefore a generalization of the classic ALS model to the case where there is cross-sectional dependence. In addition, spatial frontier models that compute efficiency using the cross-sectional specific effects assume that all the latent heterogeneity is inefficiency. The model we present distinguishes not just between latent heterogeneity and NVI, which is a common approach in the non-spatial stochastic frontier literature (Greene, 2005a; 2005b; Chen *et al.*, 2014; Tsionas, 2002), but also between latent heterogeneity and NII which has only recently been proposed in the non-spatial setting (Colombi *et al.*, 2011; Filippini and Greene, 2016). Rigidities in assets and rigidities in the internal organization of production would be sources of NII and concurrent with these rigidities managerial inefficiency would be a source of NVI. Managerial inefficiency will vary over time because of, for example, the turnover of managerial staff with different skill sets.

In our model latent heterogeneity is modeled using random effects rather than fixed effects or correlated random effects. The estimation procedure we propose relies on all the error components being i.i.d., which will not be the case with the fixed effects model if the fixed effects are correlated with the time-varying errors. As a result of this correlation the spatial Hausman test (Mutl and Pfaffermayr, 2011) will indicate that a fitted one-way SAR fixed effects model is preferred in the first step of the estimation procedure to the corresponding random effects specification. Following on from our model, development of a SAR stochastic frontier model with fixed effects is therefore the next logical contribution to the spatial frontier literature. One possible approach to the development of a spatial stochastic frontier with fixed effects would be to extend the non-spatial fixed effects Hausman-Taylor stochastic frontier estimator (Amsler and Schmidt, 2015) to the spatial case by drawing on the spatial non-frontier Hausman-Taylor model in Baltagi *et al.* (2015). In similar spirit another contribution that follows on from our model would involve extending to the spatial setting the literature on non-spatial stochastic frontier models that allow the random effects to be correlated with the regressors (e.g. Park *et al.*, 2003).

Our second extension of GKS focuses on a set of Monte Carlo experimental results that examine the finite sample performance of our SAR stochastic frontier and the corresponding non-spatial stochastic frontier. Our data generating process (DGP) is calibrated using rook contiguity based spatial interaction among the cross-sectional units to mirror the widely assumed spatial interaction in Monte Carlo studies of spatial non-frontier estimators (e.g. Baltagi *et al.*, 2003; 2007) and empirical applications (e.g. Fredriksson and Millimet, 2002). Our DGP is also based on a mix of parameter values from Monte Carlo studies of spatial non-frontier models (e.g. Pace and Barry, 1997) and non-spatial frontier estimators (e.g. Chen *et al.*, 2014).

Our third extension of GKS involves introducing the concept of a spatial efficiency multiplier which we use to obtain partitioned efficiencies across space. We can relate the partitioned efficiency spillovers to 1st order neighbors, 2nd order neighbors, etc. The partitioned efficiency spillovers indicate the speed of decay of efficiency spillovers across space. We also use the spatial efficiency multiplier to distinguish between the efficiency benchmark from the structural form of our SAR stochastic frontier model and the efficiency benchmark from the reduced form of the model. Following Anselin (2003) from the spatial non-frontier literature, the structural form of our model includes the endogenous SAR variable as a regressor that shifts the frontier technology, whereas the SAR variable does not feature in the reduced form of the model. Instead, the reduced form expression for our model contains only the exogenous variables, components of the composed disturbance and spatial transformations of the exogenous variables and error components. The upshot is that the structural form of our model accounts for SAR dependence but yields an own efficiency measure that is directly comparable to that from a non-spatial stochastic frontier and thus relates to an individual unit and does not include any efficiency spillovers across the system/network. The reduced form of our model, on the other hand, yields a system/network measure of efficiency for a productive unit that includes efficiency spillovers. Since the structural form of our model and its reduced form yield different measures of efficiency, we show that the appropriate efficiency benchmarks for the two forms of the model differ.

Putting our model into a general context, we can view own latent efficiency from the structural form of our model as a placeholder (normalized to have one-sided support) for productivity effects relating to either technical efficiency (ALS), cost efficiency (Olley and Pakes, 1996), intangible capital (Corrado *et al.*, 2009), unobserved organizational capital (Brynjolfsson and Hitt, 2003), or simply an unobservable factor (Levinsohn and Petrin, 2003). None of these literatures, however, address the issues we address in this paper. Using the structural form of our model we control for the effect of spillovers on own latent efficiency and interpret this own efficiency as a placeholder for own productivity effects having accounted for spatial interaction. Moreover, using the reduced form of our model we compute two asymmetric system/network measures of efficiency for a unit comprising own latent efficiency and asymmetric efficiency spillovers to or from the unit. These system/network measures of efficiency extend the above literatures as we interpret these measures as placeholders for system/network productivity effects.

The remainder of the paper is organized as follows. In section 2 we present the random effects SAR stochastic frontier model and set out the ML estimation procedure. Section 3 introduces the concept of a spatial efficiency multiplier and discusses how to partition efficiencies across space. We also discuss in more depth the difference between the efficiency metric from the structural form of our SAR stochastic frontier model and the metric from its reduced form. In section 4 we present the Monte Carlo results for our spatial stochastic frontier model and the corresponding non-spatial specification. In section 5 we apply our spatial model to a state level cost frontier for U.S. agriculture and section 6 concludes.

2 Random Effects SAR Stochastic Frontier Model

2.1 Structural Form of the Model

We extend the SAR stochastic frontier model for panel data in GKS by augmenting their model with random effects to account for latent heterogeneity and by distinguishing between net time-invariant inefficiency (NII) and net time-variant inefficiency (NVI). We consider a model that is linear in parameters and we focus our interpretations of the spatial panel frontier with random effects using a production function framework. Such a productivity model scenario, however, need not constrain the use of our estimator to only this type of setting. Keeping to a productivity-type interpretation of our spatial frontier model, dual frontier representations of the underlying technology could also be considered such as profit, revenue or cost functions that are linear in parameters. The general form of our model, which is the SAR counterpart of the non-spatial model in Colombi *et al.*, (2011) and Filippini and Greene (2016), is:

$$
\begin{aligned}
& y_{it} = \alpha + x_{it}'\beta + \delta \sum_{j=1}^{N} w_{ij} y_{jt} + \kappa_i + v_{it} - \eta_i - u_{it}, \qquad (1) \\
& \kappa_i \sim N(0, \sigma_\kappa^2),\ \ \eta_i \sim N^+\left(0, \sigma_\eta^2\right); \\
& v_{it} \sim N\left(0, \sigma_v^2\right),\ u_{it} \sim N^+\left(0, \sigma_u^2\right); \\
& \quad i = 1, ..., N;\ t = 1, ..., T.
\end{aligned}
$$

Here there are N units in each cross-section indexed $i = 1, ..., N$ that operate over T periods indexed $t = 1, ..., T$. In line with the spatial econometrics literature we consider the case of large N and fixed T. y_{it} is the observation for the dependent variable (output, cost, profit or revenue) for the ith unit at time period t, where lower case letters denote logged variables. α is the intercept, x_{it}' is a $(1 \times K)$ vector of observations for the exogenous regressors and β is a $(K \times 1)$ vector of regression parameters. x_{it} will include variables which together with y_{it} represent the frontier technology and x_{it} will also include any variables that shift the frontier. W_N is specified *a priori* and denotes the $(N \times N)$ spatial weights matrix of non-negative w_{ij} constants. W_N captures the spatial arrangement of the cross-sectional units and also the strength of the spatial interaction among the units. In the spatial econometrics literature, a measure of geographical proximity is frequently used to specify W_N. $\sum_{j=1}^{N} w_{ij} y_{jt}$ is the endogenous spatial lag of the dependent variable which shifts the frontier technology and δ is the SAR scalar parameter. Following Anselin (2003) in the spatial non-frontier literature, we refer to Eq. 1 as the structural form of the model as it contains the SAR variable as a regressor, whereas in the expression for the reduced form of the model that we use to introduce the concept of a spatial efficiency multiplier the SAR variable does not feature (see Eq. 14). Furthermore, Eq. 1 encompasses the spatial Durbin specification. In this case x_{it} will also include as additional regressors that shift the frontier technology observations for exogenous spatial lags of the independent variables.

Eq. 1 has a four component error structure $\varepsilon_{it}^{*} = \varepsilon_i + \varepsilon_{it} = \kappa_i + v_{it} - \eta_i - u_{it}$, where $\varepsilon_i = \kappa_i - \eta_i$ is the time-invariant component and $\varepsilon_{it} = v_{it} - u_{it}$ is the time-variant component. We make standard assumptions when modeling latent heterogeneity using a random effects error component specification and assume that v_{it} and u_{it} are i.i.d. across i and t, and κ_i and η_i are i.i.d. across i. We further assume that the unit specific effect κ_i is time-invariant, and that v_{it} is the idiosyncratic error, η_i is NII and u_{it} is NVI. Following the canonical ALS stochastic frontier model we assume that η_i and u_{it} have a half-normal distribution. Our model, however, is sufficiently general to accommodate alternative distributional possibilities for η_i and u_{it}.

Additional assumptions $1-4$ are based on standard normalizations and regularity conditions from the spatial econometrics literature (e.g. Baltagi *et al.*, 2003; 2007; Kelejian and Prucha, 2004).

Assumption 1: *The elements on the main diagonal of the non-stochastic W_N are zero.*

Assumption 2: *The matrix $(I_N - \delta W_N)$ is non-singular for all $\delta \in (1/h_{\min}, 1/h_{\max})$, where $h_{\min}$ and $h_{\max}$ are the most negative and most positive real characteristic roots of W_N and we assume W_N before normalization $\left(\widetilde{W_N}\right)$ is symmetric.*

Assumption 3: *The row and column sums of the matrix $\widetilde{W_N}$ are uniformly bounded in absolute value as $N \to \infty$, and uniformly in δ the matrix $\left(I_N - \delta \widetilde{W_N}\right)^{-1}$ is uniformly bounded in both row and column sums as $N \to \infty$.*

Assumption 4: *The $(N \times K)$ regressor matrix X_t is of full rank K, the elements of X_t are non-stochastic and are uniformly bounded in absolute value in N and T, and $\lim_{N\to\infty} (1/N) X'X$ exists and is non-singular.*

Assumption 1 is a normalization rule which ensures that no unit can be viewed as its own neighbor. Assumption 2 ensures that the reduced form of Eq. 1 exists. If $\widetilde{W_N}$ is asymmetric it may have complex roots so in this case $h_{\min}$ in Assumption 2 is the most negative pure real characteristic root of W_N. In the application section of the paper for all specifications of W_N the normalizations yield $h_{\max} = 1$, which simplifies the computation of $\log |I_N - \delta W_N|$ in the estimation (see subsection 2.2 for further details on this). As a result of Assumption 3 the spatial process of the dependent variable is limited to a manageable degree as it has a fading memory (Kelejian and Prucha, 1998; 2001). Assumption 4 rules out perfect collinearity and allows standard limit theorems to be utilized in deriving asymptotics. Moreover, by assuming the elements of X_t are non-stochastic, Assumption 4 rules out unbalanced panels which is convenient because the asymptotic properties of spatial estimators become problematic when the source of missing data is not known (Elhorst, 2009).[2]

2.2 Estimation and the Net Components of Gross Time-Variant Efficiency

The ML estimation procedure we set out for Eq. 1 can also be the basis for the development of ML procedures for other spatial stochastic frontier models. These models can be developed by extending spatial non-frontier specifications such as a spatio-temporal model (e.g. Elhorst 2001), a higher order SAR model (e.g. Badinger and Egger, 2011; Elhorst *et al.*, 2012) or a model with SAR and spatial autocorrelated error terms (e.g. Kelejian and Prucha, 1998; 2010) to a stochastic frontier setting.

Our three-step ML estimation procedure first involves reparameterizing Eq. 1 by using $\sigma^2_{\eta\kappa} = \sigma^2_\eta + \sigma^2_\kappa$ and $\lambda_{\eta\kappa} = \sigma_\eta/\sigma_\kappa$, and $\sigma^2_{uv} = \sigma^2_u + \sigma^2_v$ and $\lambda_{uv} = \sigma_u/\sigma_v$. Therefore, $\sigma^2_\eta = \sigma^2_{\eta\kappa}/\left(1 + \lambda^2_{\eta\kappa}\right)$, $\sigma^2_\kappa = \sigma^2_{\eta\kappa}\lambda^2_{\eta\kappa}/\left(1 + \lambda^2_{\eta\kappa}\right)$, $\sigma^2_u = \sigma^2_{uv}/\left(1 + \lambda^2_{uv}\right)$ and $\sigma^2_v = \sigma^2_{uv}\lambda^2_{uv}/\left(1 + \lambda^2_{uv}\right)$. We then rewrite Eq. 1 as follows by transforming the positively skewed time-invariant error, ε_i, the positively skewed time-variant error, ε_{it}, and the intercept.

$$y_{it} = \alpha^\circ + x'_{it}\beta + \delta\sum_{j=1}^{N} w_{ij}y_{jt} + \varepsilon^\circ_i + \varepsilon^\circ_{it}, \tag{2}$$

where $\alpha^\circ = \alpha - \mu_{\varepsilon_i} - \mu_{\varepsilon_{it}}$, $\varepsilon^\circ_i = \kappa_i - \eta_i + \mu_{\varepsilon_i}$, $\varepsilon^\circ_{it} = v_{it} - u_{it} + \mu_{\varepsilon_{it}}$, $\mu_{\varepsilon_{it}} = E\left(u_{it}\right)$ and $\mu_{\varepsilon_i} = E\left(\eta_i\right)$. Eq. 2 therefore has the form of a one-way panel model with time-invariant and time-variant error components that satisfy the zero-mean condition by construction.

Step 1 estimates Eq. 2 using ML. The log-likelihood function for the model represented in Eq. 2 is:

[2] Spatial estimators for unbalanced panels therefore involve making an assumption about why observations are missing. For example, Pfaffermayr (2013) allows elements of X_t to be stochastic in an unbalanced spatial panel by assuming that observations are missing at random.

$$LL = -\frac{NT}{2}\log\left(2\pi\sigma^2\right) + T\log|I_N - \delta W_N| - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[y_{it}^{\bullet} - \delta\left(\sum_{j=1}^{N} w_{ij}y_{jt}\right)^{\bullet} - x_{it}^{\bullet\prime}\beta\right]^2, \tag{3}$$

where $T\log|I_N - \delta W_N|$ represents the contribution to the log-likelihood from the Jacobian of the transformation from $\varepsilon_{it}^{\bullet}$ to $y_{it}^{\bullet}$. ML estimation of a spatial model such as Eq. 2 is standard in the spatial literature and the transformation from $\varepsilon_{it}^{\bullet}$ to $y_{it}^{\bullet}$ accounts for the endogeneity of the SAR variable (Anselin, 1988, pp. 63; Elhorst, 2009). $\bullet$ denotes a transformation of the variables which is dependent on θ, where θ denotes the weight attached to the cross-sectional component of the data and $0 < \theta^2 = \sigma_{uv}^2 / \left(T\sigma_{\eta\kappa}^2 + \sigma_{uv}^2\right) \leq 1$. This notation is used in order to express the transformed data in quasi-differenced form as:

$$y_{it}^{\bullet} = y_{it} - (1-\theta)\frac{1}{T}\sum_{t=1}^{T} y_{it}, \tag{4}$$

$$x_{it}^{\bullet} = x_{it} - (1-\theta)\frac{1}{T}\sum_{t=1}^{T} x_{it}. \tag{5}$$

If $\theta = 0$ the $\bullet$ transformation simplifies to the demeaning procedure and the model collapses to the one-way SAR fixed effects model. We therefore assume $\theta > 0$ rather than $\theta \geq 0$ to rule out the possibility of the model collapsing to the one-way SAR fixed effects model. This is because with the one-way SAR fixed effects model, there is the possibility that the fixed effects may be correlated with ε_{it}° in Eq. 2, which violates the i.i.d. condition of the error components that our estimator relies on.

To obtain the conditional expectation $E(y|\beta,\delta)$ we use the following concentrated log-likelihood function:

$$LL_C = \Upsilon - \frac{NT}{2}\log\left[\left(e_0^{\bullet} - \delta e_1^{\bullet}\right)'\left(e_0^{\bullet} - \delta e_1^{\bullet}\right)\right] + T\log|I_N - \delta W_N|, \tag{6}$$

where Υ is a constant that does not depend on δ. $e_0^{\bullet}$ and $e_1^{\bullet}$ are the OLS residuals from the regressions of $y_t^{\bullet}$ and $(I_T \otimes W_N)y_t^{\bullet}$ on $x_t^{\bullet}$, respectively, where I_T is the $(T \times T)$ identity matrix and $\otimes$ is the Kronecker product. This concentrated log-likelihood function can only be maximized numerically as a closed form solution for the unknown parameter δ does not exist. Before commencing the iterative procedure to maximize Eq. 6, we follow Pace and Barry (1997) for spatial non-frontier models and calculate $\log|I_N - \delta W_N|$ for a vector of values of δ over the interval $(1/h_{\min}, 1)$. As they suggest, we calculate $\log|I_N - \delta W_N|$ for values of δ based on 0.001 increments over the above feasible range for δ. Given the numerical estimate of δ, the estimator for β is:

$$\widehat{\beta} = b_0 - \delta b_1 = \left(x^{\bullet\prime} x^{\bullet}\right)^{-1} x^{\bullet\prime}[y^{\bullet} - \delta \left(I_T \otimes W_N\right) y^{\bullet}], \tag{7}$$

where b_0 and b_1 are the OLS estimates from the regressions of $y_t^{\bullet}$ and $(I_T \otimes W_N) y_t^{\bullet}$ on $x_t^{\bullet}$, respectively.

Following LeSage and Pace (2009) we obtain the standard errors using a mixed analytical-numerical Hessian, where all the second order derivatives are computed analytically, with the exception of $\partial^2 LL/\partial\delta^2$ which is evaluated numerically. Evaluating the second order derivatives of the log-likelihood function analytically rather than numerically is less sensitive to badly scaled data, and numerical rather than analytical evaluation of $\partial^2 LL/\partial\delta^2$ when N is very large avoids any computational difficulties associated with the evaluation of the large dense matrix $\left(I_N - \delta W_N\right)^{-1}$. Given β, δ and σ^2, the concentrated log-likelihood function we use to compute θ is:

$$LL(\theta) = -\frac{NT}{2} \log\left[e(\theta)' e(\theta)\right] + \frac{N}{2} \log \theta^2, \tag{8}$$

$$e(\theta)_{it} = y_{it} - (1-\theta) \frac{1}{T} \sum_{t=1}^{T} y_{it} - \delta \left[\sum_{j=1}^{N} w_{ij} y_{jt} - (1-\theta) \frac{1}{T} \sum_{t=1}^{T} w_{ij} y_{jt} \right] - \left[x_{it} - (1-\theta) \frac{1}{T} \sum_{t=1}^{T} x_{it} \right]' \beta. \tag{9}$$

Step 2 estimates λ_{uv} from $\varepsilon_{it}^{\circ} = v_{it} - u_{it} + \mu_{\varepsilon_{it}}$ using $\widehat{\varepsilon}_{it}^{\circ}$ from step 1. This is done by maximizing the concentrated log-likelihood function in Eq. 10 below with respect to λ_{uv}, which is adapted from the non-spatial two-step ML stochastic frontier estimator for cross-sectional data in Fan *et al.* (1996). Here we build on their estimator by extending it to the case of panel data in a spatial setting, we account for latent heterogeneity via random effects, and we add a third step to the estimation procedure to distinguish between latent heterogeneity and the time-invariant inefficiency component NII. Moreover, whereas we use parametric estimation for all three steps they use a non-parametric method for step 1 and a parametric estimator for step 2. Specifically, in step 1 they use a kernel estimator to calculate the expected value of the dependent variable and in step 2 they use ML. The new likelihood function concentrated in terms of λ_{uv} and the solution for $\widehat{\sigma}_{uv}$ are:

$$LL(\lambda_{uv}) = -NT \ln \widehat{\sigma}_{uv} + \sum_{i=1}^{N}\sum_{t=1}^{T} \ln \left[1 - \Phi\left(\frac{\widehat{\varepsilon}_{it}^{\circ}\lambda_{uv}}{\sigma_{uv}}\right)\right] - \frac{1}{2\widehat{\sigma}_{uv}^{2}} \sum_{i=1}^{N}\sum_{t=1}^{T} \widehat{\varepsilon}_{it}^{\circ 2}, \tag{10}$$

$$\widehat{\sigma}_{uv} = \left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[y_{it}^{\bullet} - \delta\left(\sum_{j=1}^{N} w_{ij}y_{jt}\right)^{\bullet} - x_{it}^{\bullet\prime}\beta\right]^{2} / \left[1 - 2\lambda_{uv}^{2}/\pi\left(1 + \lambda_{uv}^{2}\right)\right]\right)^{1/2}, \tag{11}$$

where Φ is the standard normal cumulative distribution function. The solution to maximizing $LL(\lambda_{uv})$ yields the ML estimate $\widehat{\lambda}_{uv}$. By substituting $\widehat{\lambda}_{uv}$ into Eq. 11 we obtain the ML estimate $\widehat{\sigma}_{uv}^{2}$. The consistent estimator of the constant term is then scaled by the value of $\widehat{\mu}_{\varepsilon_{it}}\left(\widehat{\lambda}_{uv}, \widehat{\sigma}_{uv}^{2}\right)$.[3]

We use the following Jondrow *et al.* (1982) method to predict time-variant technical inefficiency u_{it} conditional on ε_{it}, where we refer to u_{it} as net time-variant inefficiency, NVI_{it}:

$$\widehat{u}_{it} = E\left(u_{it}|\varepsilon_{it}\right) = \frac{\sigma_u\sigma_v}{\sigma_{uv}}\left(\frac{\phi_{it}}{1 - \Phi_{it}} - \frac{\varepsilon_{it}\lambda_{uv}}{\sigma_{uv}}\right), \tag{12}$$

where $\Phi_{it} = \Phi\left(\varepsilon_{it}\lambda_{uv}/\sigma_{uv}\right)$, $\phi_{it} = \phi\left(\varepsilon_{it}\lambda_{uv}/\sigma_{uv}\right)$, Φ is as previously defined and ϕ is the probability density function for the standardized normal.

Recall that the dependent and independent variables in Eq. 1 entered in log form. Then the multiplicative form of our new SAR stochastic frontier model is:

$$Y_{it} = \exp\left(\alpha\right) * X_{it}^{\beta} * \left(\sum_{j=1}^{N} w_{ij}Y_{jt}\right)^{\delta} * NVE_{it} * NIE_{i} * \exp\left(v_{it} + \kappa_i\right), \tag{13}$$

where Y_{it} is a scalar observation of the dependent variable (output, cost, profit or revenue) for the ith unit at time period t. X_{it} is a $(K \times 1)$ vector of scalar observations for the exogenous regressors and everything else is as previously defined. Productive units may of course lie below the concave production, profit or revenue frontier or above the convex cost frontier because they are inefficient. Lower NVI_{it} (NII_i) will push NVE_{it} (NIE_i), which is bounded in the interval $[0, 1]$, closer to the upper bound. As a result of our new SAR stochastic frontier model in Eq. 1 having the multiplicative form in Eq. 13 $NVE_{it} = \exp\left(-\widehat{u}_{it}\right)$. NII_i is computed in a similar way to NVI_{it} and involves utilizing in step 3 a similar ML estimator to that for step 2 to estimate $\lambda_{\eta\kappa}$ from $\varepsilon_{i}^{\circ} = \kappa_i - \eta_i + \mu_{\varepsilon_i}$ based on $\widehat{\varepsilon}_{i}^{\circ}$ from step 1. NIE_i is calculated in a similar way to NVE_{it} using $-\widehat{\eta}_i$ and gross time-

[3] For stochastic production, revenue and profit frontiers that have the general form of Eq. 1, a productive unit is assumed to maximize the objective variable so for these technologies the constant is scaled down by $\widehat{\mu}_{\varepsilon_{it}}\left(\widehat{\lambda}_{uv}, \widehat{\sigma}_{uv}^{2}\right)$. Conversely, for a stochastic cost frontier the constant is scaled up by $\widehat{\mu}_{\varepsilon_{it}}\left(\widehat{\lambda}_{uv}, \widehat{\sigma}_{uv}^{2}\right)$.

variant efficiency (i.e. combined efficiency) is $GVE_{it} = \exp\left[-\left(\widehat{\eta}_i + \widehat{u}_{it}\right)\right] = NIE_i{*}NVE_{it}$.

3 Spatial Efficiency Multiplier and the Reduced Form of the Model

Relating the general spatial econometrics literature (e.g., LeSage and Pace, 2009) to our model in Eq. 1, for all the productivity oriented functional forms of our model that are linear in parameters including flexible functional forms of the technology (e.g., the translog function, generalized Leontief function or quadratic function), the marginal effect for an independent variable is a function not just of the slope parameters for the exogenous regressors but also the SAR parameter. For our model we obtain the marginal effects for the independent variables and the associated standard errors by following the general spatial econometrics literature which involves computing the direct, indirect and total marginal effects. If we consider a model such as Eq. 1 that is linear in logs then these marginal effects are unit free elasticity measures. We follow this interpretation of the direct, indirect and total marginal effects throughout the remainder of the paper.

For a SAR stochastic frontier model, a direct elasticity is interpreted in the same way as the corresponding elasticity from a non-spatial frontier, although a direct elasticity takes into account feedback effects. This feedback is the effect of a change in an independent variable of a particular unit which as a result of the spatial multiplier $(I_N - \delta W_N)^{-1}$ partially rebounds back to the dependent variable of the same unit via the effect on the dependent variables of 1st and higher order neighbors. Only partial rebound is possible because of the fading memory across space (Assumption 3 above). An indirect elasticity from a SAR stochastic frontier model can be calculated in two ways yielding the same numerical value. This leads to two interpretations of an indirect elasticity: (i) average change in the dependent variable of all other units following a change in an independent variable for one particular unit; or (ii) average change in the dependent variable for a particular unit following a change in an independent variable for all the other units. The total elasticity for a SAR stochastic frontier model is the sum of the direct and indirect elasticities. To compute the direct, indirect and total elasticities we differentiate the following reduced form of Eq. 1:

$$y_t = (I_N - \delta W_N)^{-1} \left(\alpha\iota + \beta' X_t + \kappa + v_t - \eta - u_t\right), \tag{14}$$

where vectors and matrices are stacked, ι is an $(N \times 1)$ vector of ones, lower case letters denote logged variables and everything else is as previously defined for Eq. 1. We use Monte Carlo simulation of the distributions of the direct, indirect and total elasticities to compute the associated standard errors which is a widely used approach in the spatial econometrics literature. This approach involves drawing $1,000$ parameter combinations of $\left(\widehat{\delta}, \widehat{\beta}, \widehat{\sigma}^2\right)$ from the variance-covariance matrix where each parameter has a random component drawn from $N(0,1)$.

From the reduced form of the SAR stochastic frontier model in Eq. 14, GKS develop an approach to compute direct, asymmetric indirect and asymmetric total efficiencies relative to the best performing unit in the sample in each period using the SS estimator. Relative direct efficiency is interpreted in a similar way to own efficiencies from a non-spatial frontier or a structural spatial frontier such as Eq. 1. Relative direct efficiency, however, also takes into account efficiency feedback. An example of such feedback is the effect of a change in an independent variable for a particular unit which affects the dependent variables and the efficiencies of the unit's 1st and higher order neighbors, and then through the spatial multiplier matrix this effect partially reverberates back to the dependent variable and efficiency of the unit which initiated the change. Moreover, the relative indirect efficiencies can be asymmetric because the relative efficiency spillovers to and from a unit can differ in magnitude. Since the relative total efficiencies are the relative direct and relative indirect efficiencies combined, asymmetric relative indirect efficiencies result in asymmetric relative total efficiencies.

Calculating relative direct, asymmetric relative indirect and asymmetric relative total efficiencies using the SS estimator is very informative as it constructs appropriate direct, indirect and total efficiency frontiers. This is important because the relative total efficiency frontiers will differ from the efficiency frontier from the structural SAR stochastic frontier model in Eq. 1. This is because the latter frontier is the efficiency benchmark for each unit's own efficiency. In contrast, the relative total efficiency frontiers are benchmarks for a unit's efficiency across a network/system. We know that own efficiency from the structural form of the SAR frontier in Eq. 1 and relative direct efficiency from the reduced form of the SAR frontier in Eq. 14 are all measures of substantive economic performance. We do not know, however, whether the relative asymmetric indirect efficiencies which GKS propose from the reduced form of the SAR frontier represent substantive economic performance spillovers. To illustrate, all units in a sample can have high relative indirect efficiencies but economic performance spillovers can be negligible because the absolute indirect efficiencies are all small. To establish whether relative indirect efficiencies relate to substantive economic performance spillovers we propose using relative indirect efficiencies alongside absolute indirect efficiencies.

In contrast to GKS who focus on the method to compute relative direct, relative asymmetric indirect and relative asymmetric total efficiencies, we focus on the method to compute the corresponding absolute efficiencies. As we will show in the application of our model to a state level cost frontier for the agricultural sector in the U.S., if the absolute indirect efficiency spillover to (from) a unit is sufficiently large the corresponding absolute total efficiency will be greater than the own efficiency benchmark of 1 from the structural SAR stochastic frontier. This clearly demonstrates that the benchmark from the structural SAR stochastic frontier for a unit's own efficiency is not the same as the benchmark for a unit's efficiency across a network/system from the reduced form of the SAR stochastic frontier. The absolute direct, absolute asymmetric indirect and absolute asymmetric total efficiencies which we propose can be easily interpreted as they are percentages, although we do not bound these efficiencies in the interval $[0, 1]$.[4]

We now turn to the method to compute the absolute direct, absolute asymmetric indirect and absolute asymmetric total efficiencies, which is more complex for a model such as Eq. 1 than the model in GKS. This is because Eq. 1 has time-variant and time-invariant inefficiency components, whereas the model in GKS is simpler because the time-invariant inefficiency component is omitted. From the reduced form of the SAR frontier in Eq. 14 we recognize that $(I_N - \delta W_N)^{-1} \eta = NII_{To}^{Tot}$ and $(I_N - \delta W_N)^{-1} u_t = NVI_{To}^{Tot}$, where NII_{To}^{Tot} and NVI_{To}^{Tot} denote $(N \times 1)$ vectors of absolute total NII and absolute total NVI, and To denotes that the absolute inefficiency spillovers used in the calculation of these absolute total inefficiency vectors are the inefficiency spillovers which come to the ith unit from all the jth units for $i \neq j$. As a result of the multiplicative form of our model in Eq. 13 we can transform NII_{To}^{Tot} and NVI_{To}^{Tot} into efficiencies, $(I - \delta W)^{-1} \exp(-\eta) = NIE_{To}^{Tot}$ and $(I - \delta W)^{-1} \exp(-u_t) = NVE_{To}^{Tot}$, where NIE_{To}^{Tot} and NVE_{To}^{Tot} denote $(N \times 1)$ vectors of absolute total NIE and absolute total NVE. Moreover, since we have established that $\exp[-(\eta + u_t)] = NIE * NVE = GVE$ then $(I_N - \delta W_N)^{-1} \exp[-(\eta + u_t)] = GVE_{To}^{Tot}$, where GVE_{To}^{Tot} denotes the $(N \times 1)$ vector of absolute total GVE. Accordingly, GVE_{To}^{Tot} can be written in the following form. Similar expressions can be derived for NIE_{To}^{Tot} and NVE_{To}^{Tot}.

$$(I_N - \delta W_N)^{-1} \begin{pmatrix} GVE_1 \\ \vdots \\ GVE_N \end{pmatrix}_t = \begin{pmatrix} GVE_{11}^{Dir} & + & \cdots & + & GVE_{1N}^{Ind} \\ \vdots & + & \ddots & + & \vdots \\ GVE_{N1}^{Ind} & + & \cdots & + & GVE_{NN}^{Dir} \end{pmatrix}_t = \begin{pmatrix} GVE_{To,\,1}^{Tot} \\ \vdots \\ GVE_{To,\,N}^{Tot} \end{pmatrix}_t, \tag{15}$$

where GVE_{ij}^{Dir} on the main diagonal is the direct GVE of a unit, GVE_{ij}^{Ind} is the indirect GVE spillover to the ith unit from the jth unit for $i \neq j$ and $GVE_{To}^{Ind} = \sum_{j=1}^{N} GVE_{ij}^{Ind}$ is the sum of the absolute indirect GVE spillovers to the ith unit from all the jth units for $i \neq j$.

The column sums of the above components is the $(1 \times N)$ absolute total GVE vector that we denote $GVE_{From}^{Tot'} = \left(GVE_{From,\,1}^{Tot}, GVE_{From,\,2}^{Tot}, \,.\,.\,, GVE_{From,\,N}^{Tot}\right)$. $From$ denotes

[4] See GKS for details of the SS method to transform these absolute efficiencies into relative efficiencies which are bounded in the interval $[0, 1]$.

that the absolute indirect GVE spillovers used in the calculation of $GVE_{From}^{Tot'}$ are the GVE spillovers that come to the jth unit from all the ith units for $i \neq j$. $GVE_{From}^{Ind} = \sum_{i=1}^{N} GVE_{ij}^{Ind}$ is the sum of the absolute indirect GVE spillovers to the jth unit from all the ith units for $i \neq j$. GVE_{To}^{Tot} and GVE_{From}^{Tot} therefore measure a unit's absolute GVE across a system/network, where in the empirical application the system is the state level agricultural sector across the U.S.

Although in spatial Monte Carlo simulations the assumed W_N may be symmetric, in empirical applications W_N will typically be asymmetric. If W_N is asymmetric $(I_N - \delta W_N)^{-1}$ will be asymmetric resulting in $GVE_{ij}^{Ind} \neq GVE_{ji}^{Ind}$ in Eq. 15 indicating that there are asymmetric absolute indirect GVE spillovers to and from a unit. Since direct, indirect and total NIE, NVE and GVE all contain some form of efficiency spillover they represent different performance metrics to own NIE, NVE and GVE from the structural form of our model in Eq. 1. The own NIE, NVE and GVE benchmarks are not therefore the appropriate benchmarks for the corresponding direct, indirect and total NIE, NVE and GVE from the reduced form of our model in Eq. 14. As we noted above, own NIE, NVE and GVE from a non-spatial model or the structural form of our new SAR model are all bounded in the interval $[0, 1]$. The lower bound of absolute direct, indirect and total NIE, NVE and GVE from the reduced form of our SAR model will of course also be 0. Other than that absolute direct, indirect and total NIE, NVE and GVE are unbounded. Absolute direct, indirect and total NIE, NVE and GVE, however, can easily be interpreted as they are percentages because they are scaled own NIE, NVE and GVE. The magnitude of the scaling relates to the magnitude of the efficiency spillover that is included in the absolute direct, indirect and total NIE, NVE and GVE. If the magnitude of the efficiency spillover is sufficiently large, absolute direct/indirect/total NIE, NVE or GVE will be greater than 1. If this is the case the efficiency spillover has pushed the unit beyond the best practice frontier for own efficiency from the structural form of our model.

As a result of the fading memory property of $(I_N - \delta W_N)^{-1}$, efficiency spillovers to and from a unit die out across space. To examine the speed of the decay of efficiency spillovers across space we recognize that $(I_N - \delta W_N)^{-1} \exp(-\eta)$, $(I_N - \delta W_N)^{-1} \exp(-u_t)$ and $(I_N - \delta W_N)^{-1} \exp[-(\eta + u_t)]$ are spatial NIE_{To}^{Tot}, NVE_{To}^{Tot} and GVE_{To}^{Tot} multipliers. The Leontief expansion of the spatial multiplier matrix up to a pre-specified order for GVE_{To}^{Tot} is as follows, where similar expressions for spatial expansions of GVE_{From}^{Tot}, NIE_{To}^{Tot}, NIE_{From}^{Tot}, NVE_{To}^{Tot} and NVE_{From}^{Tot} can be derived.

$$(I_N - \delta W_N)^{-1} \exp[-(\eta + u_t)] = \left(I_N + \delta W_N + \delta^2 W_N^2 + \delta^3 W_N^3 + ...\right) \exp[-(\eta + u_t)]. \tag{16}$$

Using Eq. 16 we thus partition GVE_{ij}^{Dir}, GVE_{To}^{Ind} and GVE_{To}^{Tot} into own GVE (I_N) and

GVE spillovers which come to the ith unit from 1st order neighbors (δW_N), 2nd order neighbors ($\delta^2 W_N^2$), etc.

4 Monte Carlo Experiments

We next turn to a simulation study of the estimator we propose for our new SAR stochastic frontier model. To the best of our knowledge Monte Carlo experiments to evaluate the finite sample performance of a spatial stochastic frontier estimator have not appeared in the literature. Accordingly, there is much scope for wider application of the simulation methodology we develop in this section. Our DGP is based on Eq. 1 with one regressor and a row-normalized W_N based on rook contiguity, which is a common specification of W_N in empirical work (e.g. Fredriksson and Millimet, 2002) and also turns out to be our preferred specification of W_N in the empirical application in the next section. Moreover, in Monte Carlo experiments of spatial non-frontier estimators the assumed spatial interaction to populate W_N is frequently row-normalized rook contiguity (e.g. Baltagi *et al.*, 2003; 2007).

As is common in Monte Carlo experiments in the spatial literature our formulation of W_N is based on a board of dimension r yielding $N = r^2$. Our W_N is therefore characterized by connectivity of four, three and two units in the inner, border and corner fields, respectively. The sample sizes that we consider are similar to those that have been used in Monte Carlo experiments in the non-spatial stochastic frontier literature (e.g. Chen *et al.*, 2014) and are combinations of T = 5, 10, 50 and N = 9, 16, 49, 100, 225, 529.[5] For each sample size we generate 1000 Monte Carlo samples and the DGP sets $\sigma_\eta = \sigma_\kappa = 0.5$ yielding $\lambda_{\eta\kappa} = 1$. We also set $\sigma_u = \sigma_v = 0.25$, which following Chen *et al.* (2014) sets $\lambda_{uv} = 1$. κ_i and η_i are generated for $i = 1, ..., N$ from the corresponding distribution in Eq. 1 and we generate v_{it} and u_{it} for $i = 1, ..., N$ and $t = 1, ..., T$. The uniform random variable x_t is generated with support on the $[0, 1]$ interval and we again follow Chen *et al.* (2014) by setting $\beta = 1$. Following Pace and Barry (1997) in the spatial non-frontier literature we set $\delta = 0.25$ and we generate y_t based on the reduced form in Eq. 14.

The simulation results for our new SAR frontier model and the corresponding non-spatial model in the presence of SAR dependence are presented in tables 1 and 2, respectively. In these tables we report for parameter estimates and expected values of NVI and NII average bias and average mean square error (MSE) across the samples for each combination of T and N. It is evident from table 1 that the simulation results for the estimates of β and δ are favorable to our new estimator because the average biases and average MSE tend to be quite small for all combinations of T and N. From the same

[5]The sample sizes we consider here are not exactly those used in Chen *et al.* (2014) because the sample sizes for their non-spatial Monte Carlo experiments are not based on $N = r^2$.

table we can also see for the SAR model that the largest average bias and MSE for the estimates of $E(u_{it}|\varepsilon_{it})$, $E(\eta_i|\varepsilon_i)$, σ_v^2, σ_u^2, σ_κ^2 and σ_η^2 is in each case reasonably small. With regard to any patterns in the direction of the simulation results we note that, on average, our new SAR model yields rather small underestimates of σ_v^2 and σ_u^2 across the experiments for all combinations of T and N.

[Insert tables 1 and 2 about here]

Comparing table 1 with the simulation results in table 2 reveals that the intercept from the non-spatial model tends to absorb the omitted SAR dependence. This is not surprising because W_N remains unchanged across all T periods in the experiments. In particular, we find that for the estimates of α from the non-spatial model for all combinations of T and N the average bias and average MSE are much larger than we observe from our new SAR model. Moreover, it is evident from table 2 that, on average, the estimates of β from the non-spatial model are only overestimated by about 0.02 across all the experiments. From tables 1 and 2 we can see that the corresponding average bias (MSE) from our new SAR model and the non-spatial specification are of a similar order of magnitude for the estimates of σ_v^2. We also find this is the case for the estimates of σ_u^2. For our new SAR model we noted above that the largest average bias and MSE in table 1 for the estimates of $E(u_{it}|\varepsilon_{it})$, $E(\eta_i|\varepsilon_i)$, σ_v^2, σ_u^2, σ_κ^2 and σ_η^2 is in each case reasonably small. For the non-spatial model, however, we find that the largest average bias (MSE) in table 2 for the estimates of $E(u_{it}|\varepsilon_{it})$, $E(\eta_i|\varepsilon_i)$, σ_v^2, σ_u^2, σ_κ^2 and σ_η^2 is in each case relatively large when compared to the largest corresponding average bias (MSE) from the SAR model.

We now examine the statistical performance of the β estimates from the Monte Carlo experiments in more detail. We focus on the β estimates because, whereas σ_u^2 and $E(u_{it}|\varepsilon_{it})$ relate to NVI and σ_η^2 and $E(\eta_i|\varepsilon_i)$ relate to NII, a β estimate has wider implications in efficiency and productivity analysis as it affects NVI and NII as well as returns to scale and the monotonicity and curvature characteristics of the technology. In table 3 we present the average 95% confidence intervals for the β estimates for each combination of T and N as well as the average coverage probability of the confidence intervals. Our findings from table 3 are also supported by the 95% confidence intervals in figure 1. From table 3 we can see for our new SAR model that the average coverage probability is high for all sample sizes and ranges from $0.94 - 0.96$. In contrast, table 3 clearly shows for the non-spatial model that the average coverage probability tends to decline sharply as the sample size increases. The extreme cases are zero coverage probabilities for $T = 50$ and $N = 225$ and $T = 50$ and $N = 529$. Zero and small coverage probabilities for the non-spatial model are because, on average, the confidence intervals for the β estimates are narrow for the larger sample sizes that we consider with lower bounds slightly above 1. The average confidence intervals for the β estimates from our new SAR model also tend to become narrow for the larger sample sizes but as we noted above the corresponding average coverage probabilities are high for all sample sizes. This is because, on average, the lower bound of the confidence intervals is less than 1.

[Insert table 3 and figure 1 about here]

5 Application to State Agriculture in the U.S.

We now present an empirical illustration of our new estimator. The general structural form of SAR production, SAR profit and SAR revenue stochastic frontiers with random effects and NII and NVI components was set out in Eq. 1. Rather than estimate a SAR production, SAR profit or SAR revenue stochastic frontier we estimate a SAR stochastic cost frontier based on agricultural data for the contiguous states in the U.S. In particular, we estimate SAR and spatial Durbin stochastic cost frontiers (the latter is denoted SDF), which are both specifications from the general class of SAR models. The SDF specification is the SAR specification augmented with spatial lags of the exogenous variables as additional regressors that shift the frontier technology. To enable comparisons we also estimate the corresponding non-spatial stochastic cost frontier model (denoted NSF).

We adopt the flexible translog cost technology and because the SDF nests the SAR and NSF models, we present the three model specifications that we estimate using only the following SDF, where lower case letters denote logged variables:

$$c_{it} = \alpha + \tau_t + TL\left(y_{it}, p_{it}, t\right) + \sum_{j=1}^{N} w_{ij} TL\left(y_{jt}, p_{jt}, t\right) + \delta \sum_{j=1}^{N} w_{ij} c_{jt} + \kappa_i + v_{it} + \eta_i + u_{it}. \tag{17}$$

Here $TL\left(y_{it}, p_{it}, t\right)$ represents the translog cost technology and is a quadratic function in outputs, input prices and time. c_{it} is total cost for the ith state at time t, α is the intercept, τ_t is a time period effect and t is a time trend.[6] y_{it} is a vector of outputs, p_{it} is a vector of input prices and δ is the SAR parameter. The error structure $\varepsilon_{it} = \kappa_i + v_{it} + \eta_i + u_{it}$ differs from that for Eq. 1 because Eq. 17 is a stochastic cost frontier; otherwise κ_i, v_{it}, η_i and u_{it} are as defined for Eq. 1. Other settings where our model can be applied are SAR stochastic input and output distance frontiers. In contrast to single output production and single output cost functions, the characterization of input distance and output distance production technologies is in terms of multiple inputs being used to

[6] We follow the spatial decomposition of aggregate TFP growth for European countries by Glass *et al.* (2013) and capture the effects of time in Eq. 17 in two ways. We first specify a non-linear time trend to capture average technical change over the study period by including t, t^2 and interactions with t. We next include time period dummy variables to capture departures from the non-linear trend in a particular year due to, for example, common macroeconomic shocks (see, for example, Baltagi and Griffin, 1988).

produce multiple outputs. A stochastic input distance frontier adopts an input conserving approach to the measurement of a productive unit's inefficiency as represented by the unit's distance from the frontier. This distance measures the maximum amount that a productive unit's input vector can be radially contracted to produce a given output vector. This setting is akin to a SAR stochastic production frontier so the error structure for a SAR stochastic input distance frontier is as in Eq. 1 and the constant is scaled down in the estimation by $\widehat{\mu}_{\varepsilon_{it}}\left(\widehat{\lambda}_{uv},\widehat{\sigma}_{uv}^{2}\right)$. Alternatively, a stochastic output distance frontier adopts an output expanding approach to the measurement of a productive unit's inefficiency. This inefficiency represents the maximum amount that a productive unit's output vector can be radially expanded using a given input vector. This setting is akin to a SAR stochastic cost frontier so the error structure for a SAR stochastic output distance frontier is as in Eq. 17 and the constant is scaled up in the estimation by $\widehat{\mu}_{\varepsilon_{it}}\left(\widehat{\lambda}_{uv},\widehat{\sigma}_{uv}^{2}\right)$.

$\sum_{j=1}^{N} w_{ij}c_{jt}$ is the spatial lag of the dependent variable and $\sum_{j=1}^{N} w_{ij}TL\left(y_{jt},p_{jt},t\right)$ is the spatial lag of $TL\left(y_{it},p_{it},t\right)$, where w_{ij} is as defined for Eq. 1. When W_N is row-normalized W_Nt and W_Nt^2 must be omitted for reasons of perfect collinearity because $W_Nt=t$ and $W_Nt^2=t^2$. Minor differences in the specifications estimated for the NSF, SAR and SDF involve terms such as these being omitted due to perfect collinearity. In our application no variables shift the cost frontier technology in the NSF, whereas $\sum_{j=1}^{N} w_{ij}c_{jt}$ shifts the cost frontier technology in the SAR model. In the SDF $\sum_{j=1}^{N} w_{ij}TL\left(y_{jt},p_{jt},t\right)$ and $\sum_{j=1}^{N} w_{ij}c_{jt}$ shift the cost frontier technology.

5.1 Data and Spatial Weights Matrices

The publicly available data are for the 48 contiguous states in the U.S. for the period $1960-2004$. The data is drawn mainly from the U.S. Department of Agriculture's (USDA) productivity database, which has been widely used in the productivity literature (e.g. O'Donnell, 2012; 2016; Griffiths and O'Donnell, 2005; Schimmelpfennig *et al.*, 2006; Karagiannis and Mergos, 2000).[7] For an overview of the USDA productivity database see Ball *et al.* (1999).[8]

The fitted non-spatial and spatial cost frontiers are long run specifications and are based on the widely used specification employed by Ray (1982). The models are estimated using the following data and variable construction. The three outputs are measured in 000s of 1996 U.S. dollars and are the implicit quantity of livestock and products (y_1), the implicit quantity of crops (y_2) and the implicit quantity of farm related output (y_3).[9] The input price data are indices for capital services (p_1), labor services (p_2), total intermediate

[7] The study period ends in 2004 because of the discontinuation of key data sources.

[8] For a more detailed discussion of the construction of the data for the variables we use see USDA (2014).

[9] Farm related output refers to goods and services from non-agricultural activities (e.g., processing and packaging of agricultural products) and secondary activities (e.g., machine services for hire).

inputs (p_3) and land service flows (p_4).[10] The total cost relative (c) is the sum of the four factor input expenditure relatives. p_1 is the normalizing input price index for c and $p_2 - p_4$, and the data for $p_1 - p_4$ and c are relative to the value for Alabama in 1996.[11] The descriptive statistics for the data are presented in table 4.

[Insert table 4 about here]

All the specifications of W that we employ before normalization are either based on inverse distances to the centroids of the nearest $3 - 7$ neighboring states (denoted $W_{3Near} - W_{7Near}$) or contiguous states (W_{Cont}). Since all the specifications of W have an assumed cut-off ($3 - 7$ nearest neighbors) or a natural cut-off (contiguous neighbors) they contain a lot of zeros and are therefore described as sparse. With such formulations of W, partitioning a spatial efficiency multiplier across space can easily be interpreted as efficiency spillovers to (from) a state from (to) 1st order neighbors, 2nd order neighbors, etc. This is in contrast to a dense specification of W based on inverse distances between each pair of state centroids, where in this case the interpretation of partitioned efficiency spillovers across space is at best opaque. Furthermore, since all the specifications of W that we employ are based on geographical location rather than, for example, interstate trade flows of agricultural products, the spatial weights are strictly exogenous.

In total we use eleven normalized specifications of W. Six of these specifications are denoted W_{Cont}^{Row} and $W_{3Near}^{Row} - W_{7Near}^{Row}$, where the superscript denotes that the matrix is row-normalized. Using a row-normalized specification of W preserves the scaling of the data because, for example, for a particular state the SAR variable will be a weighted average of the dependent variable for the states in its neighborhood set. When an inverse distance based spatial weights matrix is row-normalized spillovers are inversely related to the relative distance between the units. On one hand this is reasonable because distance can be viewed as a relative measure which will vary from state-to-state depending on how isolated a state is from other states. On the other hand, it could be argued this is unreasonable because the information on absolute distances between states is lost by row-normalizing. To address this issue the five remaining specifications of W, which are denoted $W_{3Near}^{Eig} - W_{7Near}^{Eig}$, are normalized by the largest eigenvalue of W before normalization. This normalization does not change the proportional relationship between the spatial weights in the specification of W before normalization, so spillovers are inversely related to the absolute distance between states.

[10] Total intermediate inputs include, for example, energy and agricultural chemicals.

[11] The dataset is constructed on the basis that each state is one large farm. Thus we do not include average farm size in a state as a variable that shifts the cost frontier technology to account for farm size scale effects. This is because these scale effects are accounted for by c and y. Also, for reasons of collinearity we do not include additional covariates that shift the cost frontier technology to account for input quality (e.g., farmer education and weather covariates to account for the quality of the labor and land inputs). This is because the input prices will reflect the quality of the inputs, among other things. Moreover, the input quantities used to calculate c are quality adjusted using hedonic measures. For more details on this quality adjustment see USDA (2014).

5.2 Estimated Models

Following the spatial analysis in Pfaffermayr (2009) model selection is based on the Akaike information criterion (AIC). We also use the Schwarz/Bayesian information criterion (SIC). Of the 23 models which we estimate (NSF, and SAR and SDF models using eleven specifications of W), the AIC and SIC both point to the W_{Cont}^{Row} SDF. This emphasizes the importance of the local spatial variables that are omitted from the W_{Cont}^{Row} SAR.[12] This is apparent from the estimation results in table 5 for our preferred W_{Cont}^{Row} SDF model and the estimation results in table 6 for the corresponding NSF and W_{Cont}^{Row} SAR models, which we report to analyze the empirical implications of the different model specifications. In particular, it is evident from the fitted W_{Cont}^{Row} SDF model that a number of local spatial parameters are significant at nominal levels (e.g., those pertaining to Wy_1, Wp_2 and Wp_4). The marginal effects from the W_{Cont}^{Row} SDF and W_{Cont}^{Row} SAR models are presented in tables 7 and 8.[13]

[Insert tables $5 - 8$ about here]

The spillover elasticities from the SAR and SDF models are the indirect marginal effects which are a function of, among other things, δ (LeSage and Pace, 2009). δ itself, however, has a meaningful interpretation as it represents the degree of SAR dependence. It is apparent from tables 5 and 6 that the estimates of δ from the W_{Cont}^{Row} SDF (0.388) and the W_{Cont}^{Row} SAR (0.339) models suggest non-negligible positive SAR dependence which is significant at the 0.1% level. We find, however, that the estimate of δ from the W_{Cont}^{Row} SDF model is significantly larger than that from the W_{Cont}^{Row} SAR model at the 5% level.

From tables $6 - 8$ we can see that each of the own/direct first order time trend parameters from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are significant at the 0.1% level.[14] As we would expect, each of these time trend parameters are all negative which indicates for the sample average state cost diminution due to technical progress. We can also see from the same tables that the own/direct output and input price elasticities (evaluated at the sample mean) from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are positive and therefore satisfy the monotonicity property of the translog cost function. The own/direct output elasticities at the sample mean from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models suggest diseconomies of scale and Wald tests reveal these returns are

[12] Unlike the endogenous SAR variable which models the global spatial dependence of the dependent variable (1st order neighbor effects all the way through to $(N-1)$th order neighbor effects), the exogenous local spatial variables, which are spatial lags of the non-spatial explanatory variables, capture only 1st order neighbor effects.

[13] The estimated parameters and standard errors for the time period dummies from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are available on request. A large number of the time period dummies from these models are significant thereby justifying their inclusion to capture significant departures from the non-linear time trend.

[14] We use mean adjusted data so that the first order own/direct/indirect/total parameter estimates from the translog cost technology of the NSF, SAR and SDF models are elasticities at the sample mean.

significantly different from 1 at the 5% level or lower. Finding that the cost technology for the U.S. agricultural sector is characterized by diseconomies of scale is in line with historical evidence (e.g. Ray, 1982). Taken together these findings suggest that operating beyond the minimum efficient scale is a persistent feature of the U.S. agricultural sector. The own/direct returns to scale from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are of the order of 0.53, 0.69 and 0.74, respectively, which suggests that economies of scale are understated when SAR dependence is overlooked.

The NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models yield own/direct elasticities at the sample mean for farm related output, y_3, which are not significant and significant own/direct elasticities for the core outputs (livestock and products, y_1, and crops, y_2). We classify y_3 as non-core because it refers to goods and services from non-agricultural activities (e.g., processing and packaging of agricultural products) and secondary activities (e.g., machine services for hire). Furthermore, although all the indirect input price and indirect output elasticities at the sample mean from our preferred W_{Cont}^{Row} SDF model are positive, only the indirect y_1, p_3 and p_4 elasticities are significant. For y_2 and p_2, however, the significant direct elasticity from the W_{Cont}^{Row} SDF model dominates the corresponding insignificant indirect elasticity. Consequently, all the total input price and total output elasticities from the W_{Cont}^{Row} SDF model are significant with the exception of the total y_3 elasticity.

5.3 Cost Efficiency Results

We next discuss the efficiency results. In our discussion we emphasize how our new spatial stochastic frontier model in Eq. 14 provides a rich set of absolute direct, indirect and total (direct and indirect combined) productivity enhancing efficiency estimates. These efficiency estimates can then be bounded in the interval $[0, 1]$ using the approach set out in GKS to compute the corresponding relative efficiency measures. To different degrees, productivity spillovers across the system/network play a role in the absolute direct, indirect and total efficiency measures that we propose. As a result of these productivity spillovers, the own efficiency benchmark from the structural form of our model in Eq. 1 will differ from the relative direct, indirect and total efficiency benchmarks. Using our new absolute direct, indirect and total efficiency measures we can quantify the gaps between the own and relative efficiency benchmarks.

5.3.1 Own Net Time-Invariant, Own Net Time-Variant and Own Gross Time-Variant Cost Efficiencies

GKS do not account for latent heterogeneity via random effects and as a result their specification measures only net time-variant efficiency, NVE, and overlooks net time-invariant efficiency, NIE. Our new model address this potential misspecification. For the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models, we present in table 9 average own NIE, average own NVE and average own gross time-variant efficiency (GVE) for the sample. In addition, we also report in table 9 average own NVE, average own NIE and average own GVE scores and the corresponding efficiency rankings for selected states. The states are selected on the basis of average total agricultural output over the sample and comprise: the three states with the highest average total output (1. California, 2. Iowa and 3. Texas); a group of states with mid-ranging average total output around the median (23. Oklahoma, 24. Kentucky, 25. Mississippi and 26. Idaho); and the three states with the lowest average total output (46. Nevada, 47. New Hampshire and 48. Rhode Island).

[Insert table 9 about here]

The sample average own NIE (NVE) scores in table 9 from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are 0.999 (0.951), 0.873 (0.666) and 1.000 (0.666), respectively. This indicates that controlling for SAR dependence and, in the case of the W_{Cont}^{Row} SDF model controlling also for local spatial dependence, leads to a substantial change in the magnitude of the gap between the sample average own NIE and the corresponding NVE. For the NSF there is only a small difference between the sample average own NIE and the corresponding NVE. In contrast, the W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models yield a sample average own NIE which is much larger than the corresponding NVE. Moreover, the average own NIE and average own NVE scores for individual states indicate that modeling spatial dependence in this application is preferred to the NSF specification. This is because we would expect to observe non-negligible variation across the average own NIE scores for the states and across the states' average own NVE estimates. As is evident from table 9, the NSF results indicate very little variation across the average own NIE estimates for the states and across the states' average own NVE scores. We can see for the W_{Cont}^{Row} SAR model from table 9, however, that there is non-negligible variation across the average own NIE scores for the states and across the states' average own NVE estimates. Table 9 also indicates for the W_{Cont}^{Row} SDF model that there is a marked variation across the average own NVE scores for the states.

The presence of non-negligible average own net time-variant inefficiency, NVI, for the sample from the W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models is entirely reasonable. Overlooking SAR dependence in the NSF specification, however, leads to only negligible average own NVI for each state, which is counterintuitive. In addition, since it is highly unlikely that rigidities in agricultural assets and rigidities in the internal organization of agricultural production would persist continuously for the duration of our 45-year sample period, own net time-invariant inefficiency, NII, is absent for the W_{Cont}^{Row} SDF model. The absence of own NII for the W_{Cont}^{Row} SDF model means that in this application the model collapses to the spatial Durbin specification of the true random effects stochastic frontier which measures only own NVE (see Greene, 2005*a*; 2005*b*, for the non-spatial specification of this model). The above rigidities, however, are likely to persist for several years which, assuming everything else remains unchanged, will be captured by a period of constant own NVI spanning possibly $5-7$ years.

For the fitted models we conduct the test in Gourieroux *et al.* (1982) of the null, $\widehat{\sigma}_G^2 = 0$, against $\widehat{\sigma}_G^2 > 0$ for $G \in \{\kappa, \upsilon, \eta, u\}$. The asymptotic distribution of the test statistic is a mixture of chi-squared distributions, $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$.[15] For the fitted W_{Cont}^{Row} SAR, W_{Cont}^{Row} SDF and NSF models, the test indicates the presence of κ, υ and u as in each case we reject the null at the 1% level. Further tests of the fitted W_{Cont}^{Row} SAR, W_{Cont}^{Row} SDF and NSF models suggest the absence of η as in each case we fail to reject the null at the 10% level. For the fitted W_{Cont}^{Row} SAR, W_{Cont}^{Row} SDF and NSF models, the absence of η (NII) means own GVE is equal to own NVE, which highlights how one can adopt a general-to-specific approach to performance measurement using the inefficiency components specification in Eq. 1.

The own GVE scores are revealing because they provide a complete picture of performance as own GVE is own NVE and own NIE combined. The sample average own GVE scores from the NSF, W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are 0.951, 0.582 and 0.666, respectively. These results again suggest that modeling spatial dependence in this application is preferred to the NSF specification because we only find negligible own gross time-variant inefficiency, GVI, for the sample average state from the NSF but, as we would expect, there is non-negligible average GVI across the sample from the W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models. Furthermore, these efficiency scores suggest that, on average, the NSF substantially overestimates own GVE. Further differences and similarities between the own efficiencies from the W_{Cont}^{Row} SAR, W_{Cont}^{Row} SDF and NSF models are evident from the average own NIE, NVE and GVE scores for individual states in table 9 and from the kernel densities of the own efficiencies in figure 2.

[Insert figure 2 about here]

5.3.2 Direct, Indirect and Total Cost Efficiencies

The direct NIE, NVE and GVE estimates from Eq. 14 are the own NIE, NVE and GVE estimates from the structural form of the model in Eq. 1 plus any efficiency feedback. Efficiency feedback is the component of a unit's efficiency that passes through neighboring units and partially rebounds back to the unit via the mechanics of the spatial efficiency multiplier. Indirect NIE, NVE and GVE scores are efficiency spillovers to and from a unit, and the total NIE, NVE and GVE scores are the sum of the direct and indirect NIE, NVE and GVE estimates. Table 10 presents for the sample average

[15] Another relevant approach to testing for the presence of each component of our error structure is due to Andrews (2001). The asymptotic distribution of the test statistic he derives allows for: (i) the possibility that the parameter value under the null lies on the boundary of the parameter space; and (ii) the possible presence of a nuisance parameter under the alternative hypothesis. The asymptotic distribution of this test statistic is not a chi-squared distribution and involves semi-parametric simulation.

state direct, indirect and total GVE scores from the specification of the W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models in Eq. 14. Table 10 also presents for illustrative purposes these average efficiencies and the corresponding efficiency rankings for the same selected states we analyze in table 9.

[Insert table 10 about here]

The sample average direct NIE score from the reduced form of the W_{Cont}^{Row} SDF model is 1.04. We noted above that the sample average own NIE score from the structural form of the W_{Cont}^{Row} SDF model is 1 which indicates that the efficiency feedback component of direct NIE for the sample average state from the reduced form of the W_{Cont}^{Row} SDF is only 4%. Although this efficiency feedback component is small its presence is sufficient to push the sample average state beyond the best practice frontier from the structural form of the W_{Cont}^{Row} SDF model which served as the benchmark for own NIE. Since from the reduced form of the W_{Cont}^{Row} SDF model the sample average state's direct $NIE > 1$, it is clear that the efficiency benchmark for own NIE from the structural form of the model is not the appropriate benchmark for direct NIE.

GKS report asymmetric sample average indirect efficiencies indicating that the efficiency spillovers to and from the sample average unit differ. This asymmetry arises because the sample average indirect efficiencies are relative measures and not, as we report here, an absolute measure. The asymmetric relative indirect efficiencies in GKS are calculated by applying the SS efficiency estimator in each period. This involves computing the indirect efficiency spillover to a unit relative to the largest efficiency spillover received by a unit in each period, which is taken to be industry best practice. The relative indirect efficiency spillover from a unit is calculated in the same way with industry best practice taken to be the largest efficiency spillover leaving a unit in each period. GKS therefore report asymmetric sample average relative indirect efficiencies because the benchmarks differ in each period. This is because in each period across the sample, the magnitudes of the largest efficiency spillovers to and from a unit differ. In contrast, the sample average absolute indirect NIE, NVE and GVE scores which we report here from the reduced forms of the W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are symmetric. From the reduced form of the W_{Cont}^{Row} SDF (W_{Cont}^{Row} SAR) model, the sample average absolute indirect NIE, NVE and GVE scores are 0.59 (0.42), 0.39 (0.32) and 0.39 (0.28), respectively. This clearly indicates for the sample average state that the reduced forms of the models yield non-negligible absolute indirect NIE, NVE and GVE spillovers.

Even though the absolute indirect NIE, NVE, and GVE scores for the sample average state are symmetric, the corresponding average absolute indirect efficiency scores for individual states are asymmetric. For example, we can see from table 10 that the average absolute indirect GVE spillovers to and from California from the reduced form of the W_{Cont}^{Row} SDF model are 0.41 and 0.28. As GKS demonstrate for the sample average unit, asymmetric relative indirect efficiencies lead to asymmetric relative total efficiencies. Since the sample average absolute indirect NIE, NVE and GVE scores from the reduced forms of the W_{Cont}^{Row} SAR and W_{Cont}^{Row} SDF models are symmetric, we also obtain symmetric absolute total NIE, NVE and GVE scores for the sample average state. The absolute total NIE, NVE and GVE scores for the sample average state from the reduced form of the W_{Cont}^{Row} SDF (W_{Cont}^{Row} SAR) model are 1.63 (1.32), 1.09 (1.00) and 1.09 (0.88), respectively. This shows that non-negligible indirect NIE, NVE and GVE spillovers can locate a unit well beyond the industry best practice frontier from the structural form of the model in Eq. 1. It is therefore particularly clear when absolute total NIE, NVE and/or $GVE > 1$ for a unit, but it is also the case when absolute total NIE, NVE and/or $GVE \leq 1$ for a unit, that the efficiency benchmarks for own NIE, NVE and GVE from the structural form of the model are not the appropriate benchmarks for absolute total NIE, NVE and GVE from the reduced form of the model. This is because own NIE, NVE and GVE and the corresponding absolute total NIE, NVE and GVE are different efficiency metrics.

As was the case for the average absolute indirect NIE, NVE and GVE scores for individual states, we can see from table 10 that the average absolute total NIE, NVE and GVE scores for individual states are also asymmetric. To illustrate, it is apparent from table 10 that average absolute GVE_{To}^{Tot} and GVE_{From}^{Tot} for California from the reduced form of the W_{Cont}^{Row} SDF model are 1.11 and 0.98. For Massachusetts, average absolute GVE_{To}^{Tot} and GVE_{From}^{Tot} from the reduced form of the W_{Cont}^{Row} SDF model are 1.23 and 1.66, respectively, which is the largest asymmetry in the sample between the average absolute total GVE scores.

5.3.3 Partitioned Direct, Indirect and Total Cost Efficiencies Across Space

We now discuss the results where we partition across space the absolute direct, absolute indirect and absolute total NIE, NVE and GVE scores into own efficiencies (W^0) and efficiencies which relate to 1st order neighbor spillovers through to 4th order neighbor spillovers (W^1– W^4). Since GVE provides a complete picture of a unit's economic performance as GVE is NIE and NVE combined, in table 11 we present the partitioned absolute direct, indirect and total GVE scores from the reduced form of the W_{Cont}^{Row} SDF model for the sample average state. We also report in table 11 from the reduced form of the W_{Cont}^{Row} SDF model the corresponding average partitioned GVE scores and the associated efficiency rankings for six selected states. The six states are once again selected on the basis of average total agricultural output over the sample period and comprise: two states with high average total output (1. California and 2. Iowa); two states with mid-ranging average total output around the median (24. Kentucky and 25. Mississippi); and two states with low average total output (46. Nevada and 47. New Hampshire).

[Insert table 11 about here]

Average direct W^0 NIE, NVE and GVE scores are the own efficiency scores from the structural form of the spatial frontier model as they omit the efficiency feedback component. Average direct W^1- W^4 NIE, NVE and GVE scores are partitioned efficiency feedback components which have rebounded back to a unit from its 1st$-$4th order neighbors. By construction direct W^1 NIE, NVE and GVE scores are zero because efficiency feedback is a 2nd or higher order neighbor phenomenon. Furthermore, we find that nearly all of the average absolute direct NIE, NVE and GVE scores from the reduced forms of the W^{Row}_{Cont} SAR and W^{Row}_{Cont} SDF models are due to average absolute direct W^0 efficiencies, or in other words, due to average absolute own efficiencies. This is because the average absolute direct W^2- W^4 efficiencies are small or even zero. For example, we can see from table 11 that the average absolute direct W^2- W^4 GVE scores for California from the reduced form of the W^{Row}_{Cont} SDF model range from $0.00 - 0.02$. Based on the small average absolute direct W^2- W^4 efficiencies which we observe and the small partitioned direct elasticities for 2nd and higher order neighbors in the spatial non-frontier literature (e.g. LeSage and Pace, 2009; Autant-Bernard and LeSage, 2011), we conclude that efficiency feedback will typically be small.

Average absolute indirect NIE, NVE and GVE scores from the reduced form of a spatial frontier model measure the sum of the NIE, NVE and GVE spillovers to (from) a unit from (to) all the other units in the sample. The average partitioned absolute indirect NIE, NVE and GVE scores which we compute measure the sum of the efficiency spillovers to (from) a unit from (to) its 1st$-$4th order neighbors. Absolute indirect W^0 NIE, NVE and GVE scores are zero by construction because indirect efficiency spillovers to and from a unit are a 1st or higher order neighbor phenomenon. From the reduced forms of the W^{Row}_{Cont} SAR and W^{Row}_{Cont} SDF models the picture is much the same for the average partitioned absolute indirect and total NIE, NVE and GVE scores from Eq. 16 as we observed for the average absolute indirect and total NIE, NVE and GVE scores from Eq. 15. This is because the partitioned absolute indirect and total NIE, NVE and GVE scores for the sample average state are symmetric but the average partitioned absolute indirect and total NIE, NVE and GVE scores for individual states are asymmetric. To illustrate these ideas, we can see from table 11 from the reduced form of the W^{Row}_{Cont} SDF model that the symmetric absolute indirect and total W^1 GVE spillovers to and from the sample average state are 0.26. In contrast, table 11 reveals for the W^{Row}_{Cont} SDF model that the asymmetric average absolute indirect and total W^1 GVE spillover to (from) California, for example, is 0.27 (0.18). Moreover, we conclude that the large indirect NIE, NVE and GVE spillovers for the sample average state and individual states are primarily due to substantial indirect W^1 NIE, NVE and GVE spillovers. This is because the partitioned indirect NIE, NVE and GVE spillovers die out across space quite quickly. This is apparent from table 11 for the the W_{Cont}^{Row} SDF model where we can see that the symmetric absolute indirect W^4 GVE spillovers for the sample average state are just 0.01.

6 Concluding Remarks

Our paper extends the emerging literature on spatial stochastic frontier models in three respects. First, we account for latent heterogeneity by developing a ML estimator of the random effects SAR stochastic frontier model. Second, we propose and implement a Monte Carlo experimental methodology to analyze the finite sample behavior of our new spatial stochastic frontier estimator. Third, we introduce the concept of a spatial efficiency multiplier which we use to partition a unit's asymmetric system/network efficiencies into own efficiency and asymmetric efficiency spillovers to (from) the unit from (to) other units located at different points in space. Using the spatial efficiency multiplier we also show that the own efficiency best practice frontier from the structural form of the SAR model is not the appropriate benchmark for the asymmetric system/network efficiencies from the reduced form of the model. This is because the asymmetric system/network efficiencies include efficiency spillovers whereas these spillovers are omitted from the own efficiency metric. Moreover, from a policy perspective system/network measures of efficiency are particularly informative. For example, our model identifies, among other things, the units that transmit the largest efficiency spillovers to other units in the sample, which is useful if a policy maker wanted to select a small number of units to maximize the impact of a targeted policy e.g., subsidies to a small number of innovative firms to develop new technologies.

There are also clear extensions of the model we propose that may be considered. We have discussed but have not detailed alternative treatments of latent heterogeneity in a spatial stochastic frontier framework e.g., a specification that allows the random effects to be correlated with the regressors. The model we propose could also be extended to allow time-varying exogenous unit specific variables (e.g., a policy variable) to affect the mean of net time-variant inefficiency (see the corresponding non-spatial specifications in Battese and Coelli, 1995, and Huang and Liu, 1994), and/or one could allow time-invariant exogenous unit specific variables (e.g., a measure of a manager's innate skills) to affect the mean of net time-invariant inefficiency. In addition to, or instead of, allowing variables to affect the mean of net time-variant inefficiency and/or the mean of net time-invariant inefficiency, one could allow net time-(in)variant inefficiency and/or the idiosyncratic error to have a more general variance structure, as is the case in the non-spatial stochastic frontier models in Caudill and Ford (1993), Caudill *et al.* (1995), Hadri (1999) and Wang (2002).

References

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

Amsler, C. and P. Schmidt (2015): 'Distinguishing between heterogeneity and inefficiency in a panel data stochastic frontier model'. Mimeo.

Andrews, D. W. K. (2001): 'Testing when a parameter is on the boundary of the maintained hypothesis'. *Econometrica*, vol. 69, pp. 683–734.

Anselin, L. (1988): *Spatial Econometrics*: *Methods and Models.* Dordrecht: Kluwer.

Anselin, L. (2003): 'Spatial externalities, spatial multipliers and spatial econometrics'. *International Regional Science Review*, vol. 26, pp. 153-166.

Autant-Bernard, C. and J. P. LeSage (2011): 'Quantifying knowledge spillovers using spatial econometric models'. *Journal of Regional Science*, vol. 51, pp. 471-496.

Badinger, H. and P. Egger (2011): 'Estimation of higher-order spatial autoregressive cross-section models with heteroscedastic disturbances'. *Papers in Regional Science*, vol. 90, pp. 213-235.

Ball, V. E., F. Gollop, A. Hawke and G. Swinand (1999): 'Patterns of state productivity growth in the U.S. farm sector: Linking state and aggregate models'. *American Journal of Agricultural Economics*, vol. 81, pp. 164-179.

Baltagi, B. H., P. H. Egger and M. Kesina (2015): 'Firm-level productivity spillovers in China's chemical industry: A spatial Hausman-Taylor approach'. *Journal of Applied Econometrics*, vol. 31, pp. 214-248.

Baltagi, B. H. and J. M. Griffin (1988): 'A General Index of Technical Change'. *Journal of Political Economy*, vol. 96, pp. 20-41.

Baltagi, B. H., S. H. Song, B. C. Jung and W. Koh (2007): 'Testing for serial correlation, spatial autocorrelation and random effects using panel data'. *Journal of Econometrics*, vol. 140, pp. 5-51.

Baltagi, B. H., S. H. Song and W. Koh (2003): 'Testing panel data regression models with spatial error correlation'. *Journal of Econometrics*, vol. 117, pp. 123-150.

Brynjolfsson, E. and L. M. Hitt (2003): 'Computing productivity: Firm-level evidence'. *Review of Economics and Statistics*, vol. 85, pp. 793-808.

Battese, G. E. and T. J. Coelli (1995): 'A model for technical inefficiency effects in a stochastic frontier production function for panel data'. *Empirical Economics*, vol. 20, pp. 325-332.

Caudill, S. B. and J. M. Ford (1993): 'Biases in frontier estimation due to heteroskedasticity'. *Economics Letters*, vol. 41, pp. 17-20.

Caudill, S. B., J. M. Ford and D. M. Gropper (1995): 'Frontier estimation and firm-specific inefficiency measures in the presence of heteroskedasticity'. *Journal of Business and Economic Statistics*, vol. 13, pp. 105-111.

Chen, Y-Y., P. Schmidt and H-J. Wang (2014): 'Consistent estimation of the fixed effects stochastic frontier model'. *Journal of Econometrics,* vol. 181, pp. 65-76.

Cliff, A. D. and J. K. Ord (1973): *Spatial Autocorrelation.* London: Pion.

Cliff, A. D. and J. K. Ord (1981): *Spatial Processes, Models and Applications.* London: Pion.

Coelli, T. J., S. Perelman and E. Romano (1999): 'Accounting for environmental influences in stochastic frontier models: With application to international airlines'. *Journal of Productivity Analysis*, vol. 11, pp. 251-273.

Colombi, R., G. Martini and G. Vittadini (2011): 'A stochastic frontier model with short-run and long-run inefficiency random effects'. Department of Economics and Technology Management, Universita Di Bergamo, Italy. Mimeo.

Cornwell, C., P. Schmidt and R. C. Sickles (1990): 'Production frontiers with cross-sectional and time-series variation in efficiency levels'. *Journal of Econometrics,* vol. 46, pp. 185-200.

Corrado, C., C. Hulten and D. Sichel (2009): 'Intangible capital and economic growth'. *Review of Income and Wealth*, vol. 55, pp. 661-685.

Druska, V. and W. C. Horrace (2004): 'Generalized moments estimation for spatial panel data: Indonesian rice farming'. *American Journal of Agricultural Economics*, vol. 86, pp. 185-198.

Elhorst, J. P. (2001): 'Dynamic models in space and time'. *Geographical Analysis*, vol. 33, pp. 119-140.

Elhorst, J. P. (2009): *Spatial panel data models.* In the *Handbook of Applied Spatial Analysis,* Fischer, M. M., and A. Getis (Eds). New York: Springer.

Elhorst, J. P., D. J. Lacombe and G. Piras (2012): 'On model specification and parameter space definitions in higher order spatial econometric models'. *Regional Science and Urban Economics*, vol. 42, pp. 211-220.

Fan, Y., Q. Li and A. Weersink (1996): 'Semiparametric estimation of stochastic production frontier models'. *Journal of Business and Economic Statistics,* vol. 14, pp. 460-468.

Filippini, M. and W. Greene (2016): 'Persistent and transient productive inefficiency: A maximum simulated likelihood approach'. *Journal of Productivity Analysis,* vol. 45, pp. 187-196.

Fredriksson, P. G. and D. L. Millimet (2002): 'Strategic interaction and the determination of environmental policy across U.S. states'. *Journal of Urban Economics*, vol. 51, pp. 101-122.

Glass, A. J., K. Kenjegalieva and J. Paez-Farrell (2013): 'Productivity growth decomposition using a spatial autoregressive frontier model'. *Economics Letters*, vol. 119, pp. 291-295.

Glass, A. J., K. Kenjegalieva and R. C. Sickles (2016): 'A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers'. *Journal of Econometrics*, vol. 190, pp. 289-300.

Gouriéroux, C., A. Holly and A. Monfort (1982): 'Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters'. *Econometrica*, vol. 50, pp. 63-80.

Greene, W. H. (2005a): 'Reconsidering heterogeneity in panel data estimators of the stochastic frontier model'. *Journal of Econometrics*, vol. 126, pp. 269-303.

Greene, W. H. (2005b): 'Fixed and random effects in stochastic frontier models'. *Journal of Productivity Analysis*, vol. 23, pp. 7-32.

Griffiths, W. and C. J. O'Donnell (2005): 'Estimating variable returns to scale production frontiers with alternative stochastic assumptions'. *Journal of Econometrics*, vol. 126, pp. 385-409.

Hadri, K. (1999): 'Estimation of a doubly heteroskedastic stochastic frontier cost function'. *Journal of Business and Economic Statistics*, vol. 17, pp. 359-363.

Huang, C. J. and J. T. Liu (1994): 'Estimation of a non-neutral stochastic frontier production function'. *Journal of Productivity Analysis*, vol. 5, pp. 171-180.

Jondrow, J., C. A. K. Lovell, I. S. Materov and P. Schmidt (1982): 'On the estimation of technical inefficiency in the stochastic frontier production function model'. *Journal of Econometrics*, vol. 19, pp. 233-238.

Karagiannis, G. and G. J. Mergos (2000): 'Total factor productivity growth and technical change in a profit function framework'. *Journal of Productivity Analysis*, vol. 14, pp. 31-51.

Kelejian, H. H. and I. R. Prucha (1998): 'A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances'. *Journal of Real Estate Finance and Economics*, vol. 17, pp. 99-121.

Kelejian, H. H. and I. R. Prucha (1999): 'A generalized moments estimator for the autoregressive parameter in a spatial model '. *International Economic Review*, vol. 40, pp. 509-533.

Kelejian, H. H. and I. R. Prucha (2001): 'On the asymptotic distribution of the Moran's I test statistic with applications'. *Journal of Econometrics*, vol. 104, pp. 219-257.

Kelejian, H. H. and I. R. Prucha (2004): 'Estimation of simultaneous systems of spatially interrelated cross sectional equations'. *Journal of Econometrics*, vol. 118, pp. 27-50.

Kelejian, H. H. and I. R. Prucha (2010): 'Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances'. *Journal of Econometrics*, vol. 157, pp. 53-67.

Levinsohn, J. and A. Petrin (2003): 'Estimating production functions using inputs to control for unobservables'. *Review of Economic Studies*, vol. 70, pp. 317-341.

LeSage, J. and R. K. Pace (2009): *Introduction to Spatial Econometrics.* Boca Raton, Florida: CRC Press, Taylor and Francis Group.

Mutl, J. and M. Pfaffermayr (2011): 'The Hausman test in a Cliff and Ord panel model'. *Econometrics Journal*, vol. 14, pp. 48-76.

O'Donnell, C. J. (2012): 'Nonparametric estimates of the components of productivity and profitability change in U.S. agriculture'. *American Journal of Agricultural Economics*, vol. 94, pp. 873-890.

O'Donnell, C. J. (2016): 'Using information about technologies, markets and firm behaviour to decompose a proper productivity index'. *Journal of Econometrics*, vol. 190, pp. 328-340.

Olley, G. S. and A. Pakes (1996): 'The dynamics of productivity in the telecommunications equipment industry'. *Econometrica*, vol. 64, pp. 1263-1297.

Pace, R. K. and R. Barry (1997): 'Quick computation of spatial autoregressive estimators'. *Geographical Analysis*, vol. 29, pp. 232-247.

Park, B. U., R. C. Sickles and L. Simar (2003): 'Semiparametric-efficient estimation of AR (1) panel data models'. *Journal of Econometrics*, vol. 117, pp. 279-309.

Pfaffermayr, M. (2009): 'Conditional $\beta-$ and $\sigma-$convergence in space: A maximum likelihood approach'. *Regional Science and Urban Economics*, vol. 39, pp. 63-78.

PFAFFERMAYR, M. (2013): 'The Cliff and Ord test for spatial correlation of the disturbances in unbalanced panel models'. *International Regional Science Review*, vol. 36, pp. 492-506.

RAY, S. (1982): 'A translog cost function analysis of U.S. agriculture, 1939-77'. *American Journal of Agricultural Economics*, vol. 64, pp. 490-498.

SCHIMMELPFENNIG, D. E., C. J. O'DONNELL AND G. W. NORTON (2006): 'Efficiency effects of agricultural economics research in the United States'. *Agricultural Economics*, vol. 34, pp. 273-280.

SCHMIDT, P. AND R. C. SICKLES (1984): 'Production frontiers and panel data'. *Journal of Business and Economic Statistics*, vol. 2, pp. 367-374.

UNITED STATES DEPARTMENT OF AGRICULTURE, USDA (2014): *Agricultural Productivity in the U.S.: Findings, Documentation and Methods*. Last accessed on 04/28/2014. Access via:

http://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us/findings,-documentation,-and-methods.aspx

TSIONAS, E. G. (2002): 'Stochastic frontier models with random coefficients'. *Journal of Applied Econometrics*, vol. 17, pp. 127-147.

WANG, H-J. (2002): 'Heteroskedasticity and non-monotonic efficiency effects of a stochastic frontier model'. *Journal of Productivity Analysis*, vol. 18, pp. 241-253.

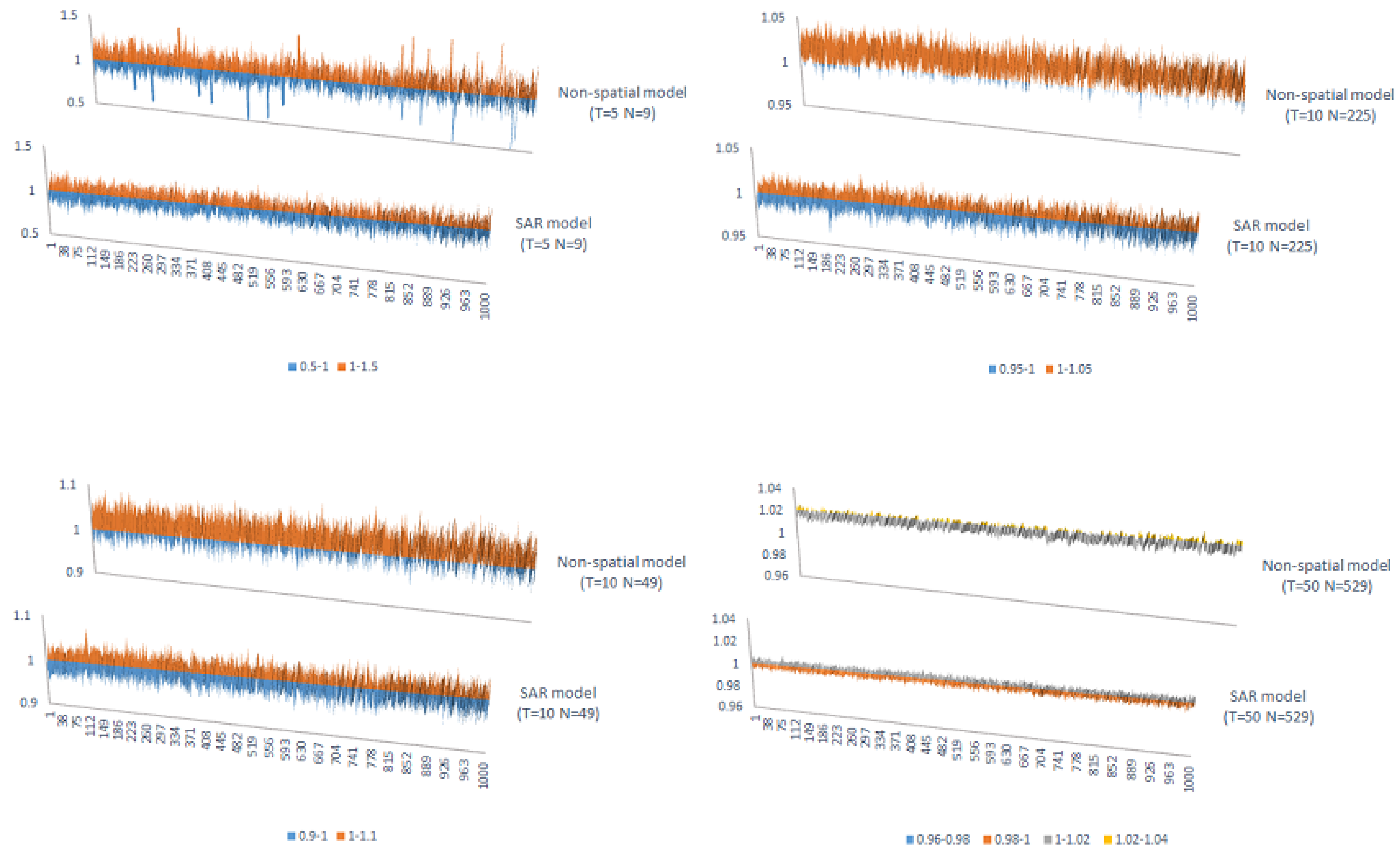

Figure 1: 95 per cent confidence intervals for the beta estimates across the 1000 Monte Carlo simulations for selected sample sizes

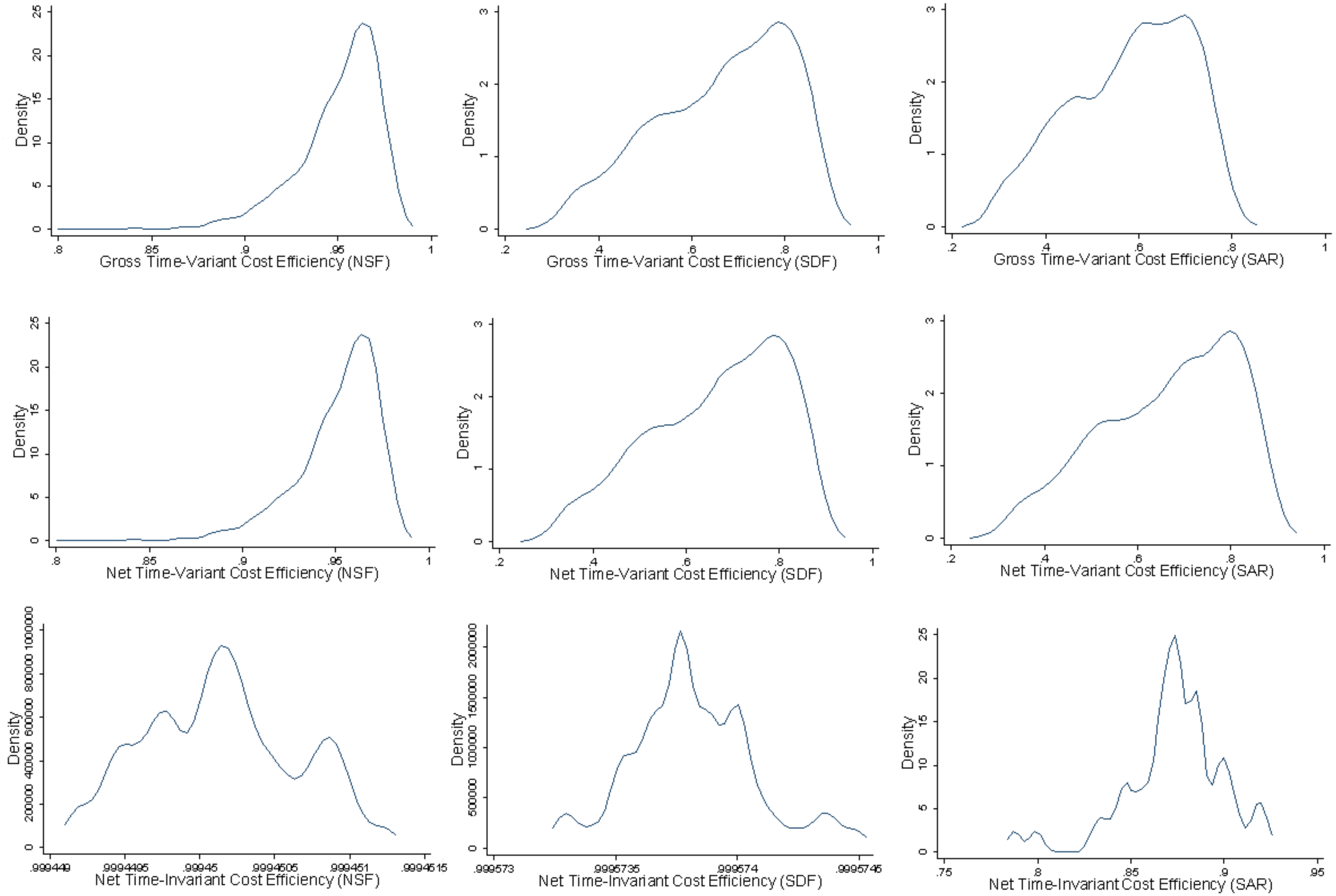

Figure 2: Kernel densities of the net time-variant, net time-invariant and gross time-variant cost efficiencies

Table 1: Performance of the SAR estimates in the presence of SAR dependence across the 1000 Monte Carlo simulations for various sample sizes

T	N		$\beta=1$	α	$\delta=0.25$	$\sigma_v^2=0.25^2$	$\sigma_u^2=0.25^2$	$E(u_{it}\|\varepsilon_{it})$	$\sigma_\kappa^2=0.5^2$	$\sigma_\eta^2=0.5^2$	$E(\eta_i\|\varepsilon_i)$
5	9	Bias	0.0004	0.1399	-0.0320	-0.0291	-0.0296	-0.0358	0.0435	-0.0641	-0.0430
		MSE	0.0026	0.0826	0.0068	0.0032	0.0041	0.0021	0.0208	0.0160	0.0074
10	9	Bias	-0.0001	0.1632	-0.0118	-0.0279	-0.0262	-0.0124	-0.0541	-0.0254	-0.0290
		MSE	0.0010	0.0476	0.0019	0.0041	0.0034	0.0018	0.0586	0.0026	0.0034
50	9	Bias	-0.0004	0.0949	0.0055	-0.0196	-0.0193	-0.0120	-0.0405	-0.0279	-0.0397
		MSE	0.0001	0.0120	0.0003	0.0011	0.0011	0.0015	0.0193	0.0030	0.0063
5	16	Bias	-0.0026	-0.0373	0.0151	-0.0193	-0.0206	-0.0231	-0.0422	-0.0292	-0.0132
		MSE	0.0009	0.0169	0.0016	0.0011	0.0013	0.0013	0.0201	0.0033	0.0007
10	16	Bias	0.0006	0.0439	-0.0003	-0.0189	-0.0111	-0.0194	0.0297	-0.0658	-0.0402
		MSE	0.0004	0.0108	0.0008	0.0011	0.0012	0.0011	0.0249	0.0160	0.0065
50	16	Bias	0.0006	-0.1151	-0.0035	-0.0189	-0.0183	-0.0079	-0.0057	-0.0663	-0.0421
		MSE	0.0001	0.0151	0.0002	0.0011	0.0010	0.0002	0.0373	0.0160	0.0071
5	49	Bias	-0.0021	0.0054	0.0152	-0.0189	-0.0184	-0.0074	0.0378	-0.0636	-0.0376
		MSE	0.0003	0.0085	0.0010	0.0011	0.0011	0.0001	0.0224	0.0158	0.0058
10	49	Bias	-0.0003	0.0122	0.0031	-0.0183	-0.0186	-0.0031	0.0254	-0.0679	-0.0423
		MSE	0.0001	0.0036	0.0003	0.0011	0.0011	0.0003	0.0263	0.0160	0.0072
50	49	Bias	-0.0001	0.0913	-0.0006	-0.0181	-0.0182	-0.0016	0.0445	-0.0646	-0.0470
		MSE	0.0000	0.0087	0.0001	0.0011	0.0011	0.0000	0.0205	0.0160	0.0088
5	100	Bias	-0.0008	-0.0227	0.0014	-0.0183	-0.0179	-0.0024	0.0290	-0.0684	-0.0425
		MSE	0.0001	0.0043	0.0004	0.0011	0.0010	0.0006	0.0251	0.0160	0.0072
10	100	Bias	-0.0001	0.0421	-0.0045	-0.0175	-0.0186	-0.0014	0.0054	-0.0631	-0.0411
		MSE	0.0001	0.0035	0.0002	0.0010	0.0011	0.0000	0.0332	0.0160	0.0068
50	100	Bias	0.0001	-0.0265	-0.0020	-0.0175	-0.0176	-0.0012	-0.0208	-0.0389	-0.0096
		MSE	0.0000	0.0011	0.0000	0.0010	0.0010	0.0000	0.0434	0.0059	0.0004
5	225	Bias	-0.0007	-0.0151	0.0078	-0.0171	-0.0183	-0.0042	0.0477	-0.0651	-0.0504
		MSE	0.0001	0.0024	0.0003	0.0010	0.0011	0.0010	0.0195	0.0160	0.0081
10	225	Bias	-0.0001	-0.0204	-0.0007	-0.0171	-0.0174	-0.0038	-0.0193	-0.0407	-0.0088
		MSE	0.0000	0.0014	0.0001	0.0010	0.0010	0.0001	0.0428	0.0065	0.0003
50	225	Bias	0.0000	0.0323	-0.0018	-0.0168	-0.0168	-0.0019	-0.0136	-0.0421	-0.0086
		MSE	0.0000	0.0011	0.0000	0.0010	0.0010	0.0000	0.0405	0.0069	0.0003
5	529	Bias	0.0000	-0.0033	-0.0027	-0.0168	-0.0171	-0.0049	-0.0072	-0.0410	-0.0094
		MSE	0.0000	0.0009	0.0001	0.0010	0.0010	0.0001	0.0379	0.0065	0.0004
10	529	Bias	-0.0002	0.0009	0.0017	-0.0165	-0.0168	-0.0032	0.0149	-0.0549	-0.0108
		MSE	0.0000	0.0004	0.0000	0.0010	0.0010	0.0001	0.0298	0.0118	0.0005
50	529	Bias	-0.0002	0.0000	0.0010	-0.0163	-0.0164	-0.0017	0.0060	-0.0495	-0.0026
		MSE	0.0000	0.0001	0.0000	0.0009	0.0010	0.0000	0.0330	0.0096	0.0000

MSE denotes Mean Square Error

Table 2: Performance of the non-spatial estimates in the presence of SAR dependence across the 1000 Monte Carlo simulations for various sample sizes

T	N		$\beta = 1$	α	$\sigma_v^2 = 0.25^2$	$\sigma_u^2 = 0.25^2$	$E\left(u_{it}\vert\varepsilon_{it}\right)$	$\sigma_\kappa^2 = 0.5^2$	$\sigma_\eta^2 = 0.5^2$	$E(\eta_i\vert\varepsilon_i)$
5	9	Bias	0.0241	1.8281	-0.0268	-0.0264	-0.0531	-0.1354	-0.1042	-0.0718
		MSE	0.0087	0.7385	0.0015	0.0015	0.0028	0.0183	0.0260	0.0124
10	9	Bias	0.0239	1.9544	-0.0261	-0.0287	-0.0560	-0.2454	-0.0371	-0.0495
		MSE	0.0022	0.9697	0.0014	0.0017	0.0036	0.0602	0.0043	0.0082
50	9	Bias	0.0245	1.9206	-0.0296	-0.0295	-0.0165	-0.2356	-0.1042	-0.0656
		MSE	0.0009	0.9038	0.0015	0.0016	0.0014	0.0584	0.0055	0.0103
5	16	Bias	0.0211	1.7979	-0.0272	-0.0287	0.0410	-0.2426	-0.0403	-0.0268
		MSE	0.0024	0.6860	0.0016	0.0017	0.0017	0.0602	0.0040	0.0018
10	16	Bias	0.0229	1.8292	-0.0268	-0.0284	-0.0312	-0.1496	-0.1042	-0.0670
		MSE	0.0013	0.7384	0.0015	0.0017	0.0011	0.0224	0.0260	0.0108
50	16	Bias	0.0221	1.6050	-0.0267	-0.0269	-0.0109	-0.1882	-0.1036	-0.0648
		MSE	0.0006	0.4033	0.0015	0.0015	0.0001	0.0354	0.0257	0.0104
5	49	Bias	0.0181	1.8562	-0.0274	-0.0272	-0.0048	-0.1602	-0.1042	-0.0711
		MSE	0.0009	0.7856	0.0016	0.0016	0.0000	0.0257	0.0260	0.0121
10	49	Bias	0.0186	1.8127	-0.0267	-0.0279	-0.029	-0.1728	-0.1042	-0.0705
		MSE	0.0006	0.7102	0.0015	0.0017	0.0006	0.0299	0.0260	0.0119
50	49	Bias	0.0190	1.9013	-0.0267	-0.0277	-0.0025	-0.1588	-0.1042	-0.0783
		MSE	0.0004	0.8673	0.0015	0.0016	0.0001	0.0252	0.0260	0.0147
5	100	Bias	0.0174	1.7578	-0.0275	-0.0277	-0.0057	-0.1639	-0.1018	-0.0575
		MSE	0.0006	0.6208	0.0016	0.0016	0.0003	0.0269	0.0249	0.0087
10	100	Bias	0.0183	1.8178	-0.0265	-0.0287	0.0029	-0.1830	-0.1042	-0.0685
		MSE	0.0005	0.7189	0.0015	0.0018	0.0004	0.0335	0.0260	0.0113
50	100	Bias	0.0181	1.7371	-0.0267	-0.0276	-0.0016	-0.2094	-0.0702	-0.0067
		MSE	0.0003	0.5884	0.0015	0.0016	0.0000	0.0438	0.0118	0.0001
5	225	Bias	0.0173	1.7965	-0.0267	-0.0288	-0.0052	-0.1537	-0.1042	-0.0842
		MSE	0.0004	0.6831	0.0015	0.0018	0.0003	0.0236	0.0260	0.0170
10	225	Bias	0.0175	1.7515	-0.0267	-0.0280	-0.0036	-0.2082	-0.0631	-0.0184
		MSE	0.0004	0.6108	0.0015	0.0017	0.0002	0.0433	0.0096	-0.0008
50	225	Bias	0.0175	1.8182	-0.0266	-0.0273	-0.0015	-0.2016	-0.0668	-0.0165
		MSE	0.0003	0.7194	0.0015	0.0016	0.0000	0.0407	0.0107	0.0007
5	529	Bias	0.0170	1.7662	-0.0270	-0.0280	-0.0046	-0.1959	-0.0624	-0.0184
		MSE	0.0003	0.6340	0.0016	0.0017	0.0002	0.0384	0.0093	0.0012
10	529	Bias	0.0170	1.7911	-0.0267	-0.0279	-0.0028	-0.1772	-0.0860	-0.0114
		MSE	0.0003	0.6743	0.0015	0.0017	0.0001	0.0314	0.0177	0.0003
50	529	Bias	0.0170	1.7868	-0.0265	-0.0275	-0.0029	-0.1823	-0.0969	-0.0045
		MSE	0.0003	0.6672	0.0015	0.0016	0.0000	0.0332	0.0225	0.0001

MSE denotes Mean Square Error

Table 3: Performance of the beta estimates across the 1000 Monte Carlo simulations for various sample sizes

		SAR Model (table 1)				Non-Spatial Model (table 2)			
T	N	Mean β (DGP $\beta = 1$)	Average Lower Bound of the 95% CI	Average Upper Bound of the 95% CI	Observed Coverage	Mean β (DGP $\beta = 1$)	Average Lower Bound of the 95% CI	Average Upper Bound of the 95% CI	Observed Coverage
5	9	1.000	0.900	1.101	0.946	1.024	0.914	1.134	0.869
10	9	1.000	0.938	1.062	0.941	1.024	0.952	1.096	0.883
50	9	1.000	0.976	1.023	0.960	1.024	0.995	1.054	0.625
5	16	0.997	0.939	1.055	0.938	1.021	0.950	1.092	0.884
10	16	1.001	0.960	1.041	0.941	1.023	0.974	1.071	0.803
50	16	1.001	0.982	1.019	0.954	1.022	1.000	1.044	0.495
5	49	0.998	0.962	1.033	0.948	1.018	0.976	1.060	0.837
10	49	1.000	0.975	1.024	0.955	1.019	0.990	1.047	0.722
50	49	1.000	0.990	1.010	0.951	1.019	1.007	1.031	0.136
5	100	0.999	0.976	1.023	0.959	1.017	0.990	1.045	0.745
10	100	1.000	0.983	1.016	0.938	1.018	0.999	1.037	0.550
50	100	1.000	0.993	1.007	0.951	1.018	1.010	1.026	0.013
5	225	0.999	0.983	1.016	0.948	1.017	0.998	1.037	0.599
10	225	1.000	0.989	1.011	0.945	1.017	1.005	1.030	0.255
50	225	1.000	0.995	1.005	0.957	1.018	1.012	1.023	0.000
5	529	1.000	0.989	1.011	0.952	1.017	1.005	1.029	0.255
10	529	1.000	0.993	1.007	0.951	1.017	1.009	1.025	0.027
50	529	1.000	0.997	1.003	0.939	1.017	1.013	1.021	0.000

Table 4: Descriptive statistics

	Variable	Mean	St. Dev.	Min	Max
Total cost relative (relative to the value for Alabama in 1996)	c	$3,105,805.49$	$3,498,059.11$	$23,561.12$	$24,027,241.07$
Livestock and products output implicit quantity (000s of 1996 U.S. dollars)	y_1	$1,677,446.55$	$1,588,017.95$	$9,100.68$	$8,497,604.24$
Crop output implicit quantity (000s of 1996 U.S. dollars)	y_2	$1,940,240.71$	$2,292,047.29$	$21,671.75$	$19,386,468.33$
Farm related output implicit quantity (000s of 1996 U.S. dollars)	y_3	$202,458.76$	$268,807.66$	798.60	$2,660,367.45$
Capital services input price index excluding land (Alabama in 1996 is 1)	p_1	0.64	0.37	0.13	1.24
Labor services input price index (Alabama in 1996 is 1)	p_2	0.44	0.33	0.05	2.11
Total intermediate inputs price index (Alabama in 1996 is 1)	p_3	0.89	0.38	0.22	2.02
Land service flows input price index (Alabama in 1996 is 1)	p_4	0.61	0.58	0.01	3.63

Table 5: Estimation results for the preferred spatial Durbin stochastic cost frontier model

				W_{Cont}^{Row} SDF				
	Coeff	t-stat		Coeff	t-stat		Coeff	t-stat
y_1	0.402***	4.73	y_3p_1	0.032**	2.64	$W(p_2p_4)$	−0.052*	−2.53
y_2	0.782***	7.46	y_3p_3	−0.056*	−2.03	$W(p_3p_4)$	0.019	0.42
y_3	0.028	0.46	y_3p_4	−0.052***	−5.67	$W(y_1y_2)$	−0.150***	−6.00
p_2	0.190***	26.43	t	−0.011***	−10.84	$W(y_1y_3)$	−0.033	−1.67
p_3	0.419***	13.81	t^2	0.000	1.57	$W(y_2y_3)$	0.055**	2.87
p_4	0.043***	3.81	y_1t	−0.001	−1.79	$W(y_1p_2)$	−0.020	−0.98
y_1^2	0.042***	5.92	y_2t	−0.002***	−3.50	$W(y_1p_3)$	−0.080	−1.87
y_2^2	0.006	0.95	y_3t	0.004***	6.05	$W(y_1p_4)$	−0.026	−1.68
y_3^2	−0.025***	−4.42	p_2t	−0.001	−0.82	$W(y_2p_2)$	0.090***	3.78
p_2^2	0.022	1.66	p_3t	0.013***	4.59	$W(y_2p_3)$	−0.033	−0.66
p_3^2	0.238***	3.99	p_4t	0.000	−0.12	$W(y_2p_4)$	0.019	1.13
p_4^2	−0.020***	−3.49	Wy_1	0.652***	3.92	$W(y_3p_1)$	−0.061**	−2.63
p_2p_3	−0.013	−0.35	Wy_2	−0.077	−0.37	$W(y_3p_3)$	0.088	1.85
p_2p_4	0.049***	4.62	Wy_3	0.114	0.99	$W(y_3p_4)$	0.023	1.32
p_3p_4	−0.099***	−3.79	Wp_2	−0.057***	−4.05	$W(y_1t)$	0.003*	2.72
y_1y_2	−0.093***	−8.41	Wp_3	−0.020	−0.39	$W(y_2t)$	−0.004**	−3.11
y_1y_3	0.002	0.18	Wp_4	0.062***	3.96	$W(y_3t)$	0.003**	2.83
y_2y_3	0.042***	4.44	Wy_1^2	0.059***	3.86	$W(p_2t)$	0.008***	4.00
y_1p_2	−0.038***	−4.00	Wy_2^2	0.054***	3.83	$W(p_3t)$	−0.001	−0.21
y_1p_3	0.049*	2.30	Wy_3^2	−0.019	−1.57	$W(p_4t)$	0.000	0.35
y_1p_4	0.009	1.24	Wp_2^2	−0.161***	−5.79	$Constant$	11.003***	6.01
y_2p_2	0.004	0.37	Wp_3^2	−0.143	−1.33	Wc	0.388***	16.05
y_2p_3	0.050	1.82	Wp_4^2	−0.048***	−4.38	$\log_{10}\theta$	−3.952***	−35.13
y_2p_4	0.054***	5.92	$W(p_2p_3)$	0.185*	2.34			
σ_v	0.2791	σ_u	0.5137		σ_κ	0.4127	σ_η	0.0005
σ_{uv}^2	0.3418	λ_{uv}	1.8408		$\sigma_{\eta\kappa}^2$	0.1703	$\lambda_{\eta\kappa}$	0.0013
LL				3078.57				
AIC				−5931.1				
BIC				−5289.5				

SDF denotes the spatial Durbin stochastic frontier model.
*, ** and *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

Table 6: Estimated non-spatial and spatial autoregressive stochastic cost frontier models

	NSF					W_{Cont}^{Row} SAR					
	Coeff	t-stat		Coeff	t-stat		Coeff	t-stat		Coeff	t-stat
y_1	0.518***	5.67	y_1p_3	0.021	0.89	y_1	0.452***	5.46	y_1p_3	0.019	0.87
y_2	1.268***	11.40	y_1p_4	0.005	0.53	y_2	0.900***	8.79	y_1p_4	0.014	1.73
y_3	0.106	1.49	y_2p_2	0.006	0.49	y_3	0.033	0.50	y_2p_2	−0.003	−0.25
p_2	0.208***	25.48	y_2p_3	0.095***	3.27	p_2	0.199***	26.97	y_2p_3	0.075**	2.86
p_3	0.509***	16.95	y_2p_4	0.068***	6.55	p_3	0.448***	16.39	y_2p_4	0.046***	4.88
p_4	0.036***	3.97	y_3p_1	0.019	1.36	p_4	0.052***	6.29	y_3p_1	0.028*	2.23
y_1^2	0.032***	3.97	y_3p_3	−0.068*	−2.29	y_1^2	0.034***	4.60	y_3p_3	−0.062*	−2.30
y_2^2	−0.003	−0.38	y_3p_4	−0.050***	−4.97	y_2^2	0.011	1.59	y_3p_4	−0.051***	−5.62
y_3^2	−0.033***	−4.93	t	−0.012***	−16.23	y_3^2	−0.033***	−5.42	t	−0.010***	−15.29
p_2^2	0.005	0.31	t^2	0.000	0.79	p_2^2	0.038**	2.66	t^2	0.000	−0.43
p_3^2	0.338***	5.43	y_1t	−0.001	−1.79	p_3^2	0.303***	5.38	y_1t	−0.001	−1.41
p_4^2	−0.044***	−7.13	y_2t	−0.001	−1.60	p_4^2	−0.031***	−5.39	y_2t	−0.001	−1.08
p_2p_3	0.025	0.54	y_3t	0.004***	6.05	p_2p_3	−0.040	−0.98	y_3t	0.003***	4.68
p_2p_4	0.061***	4.96	p_2t	0.001	0.60	p_2p_4	0.056***	5.03	p_2t	−0.001	−0.72
p_3p_4	−0.181***	−6.84	p_3t	0.017***	5.72	p_3p_4	−0.135***	−5.60	p_3t	0.016***	6.01
y_1y_2	−0.095***	−7.40	p_4t	−0.003***	−3.88	y_1y_2	−0.100***	−8.61	p_4t	−0.002**	−3.03
y_1y_3	0.017	1.62	$Constant$	8.790	9.18	y_1y_3	0.023	2.42	$Constant$	5.915***	6.71
y_2y_3	0.035**	3.09	θ	0.967***	133.22	y_2y_3	0.034***	3.30	Wc	0.339***	20.14
y_1p_2	−0.038***	−3.54				y_1p_2	−0.033***	−3.33	$\log_{10}\theta$	−3.650***	−32.32
σ_v	0.0507		σ_κ	0.3455		σ_v	0.2774		σ_κ	0.3253	
σ_u	0.0624		σ_η	0.0007		σ_u	0.5138		σ_η	0.1646	
σ_{uv}^2	0.0065		$\sigma_{\eta\kappa}^2$	0.1194		σ_{uv}^2	0.3409		$\sigma_{\eta\kappa}^2$	0.1329	
λ_{uv}	1.2303		$\lambda_{\eta\kappa}$	0.0020		λ_{uv}	1.8523		$\lambda_{\eta\kappa}$	0.5059	
LL			2705.09			LL			2883.87		
AIC			−5252.2			AIC			−5607.7		
BIC			−4803.6			BIC			−5153.5		

NSF denotes the non-spatial stochastic frontier model and SAR denotes the spatial autoregressive stochastic frontier model.
*, ** and *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

Table 7: Marginal effects from the preferred spatial Durbin stochastic cost frontier model

		SDF W_{Cont}^{Row}					
		Marginal Eff.	t-stat			Marginal Eff.	t-stat
y_1	*Direct*	0.490***	6.35	y_1p_2	*Direct*	−0.042***	−4.23
	Indirect	1.223***	4.58		*Indirect*	−0.049	−1.55
	Total	1.713***	5.63		*Total*	−0.090*	−2.54
y_2	*Direct*	0.815***	6.48	y_1p_3	*Direct*	0.041*	1.97
	Indirect	0.335	1.05		*Indirect*	−0.095	−1.58
	Total	1.149**	2.91		*Total*	−0.054	−0.79
y_3	*Direct*	0.045	0.68	y_1p_4	*Direct*	0.007	0.97
	Indirect	0.202	1.24		*Indirect*	−0.036	−1.57
	Total	0.248	1.30		*Total*	−0.029	−1.13
p_2	*Direct*	0.192***	26.58	y_2p_2	*Direct*	0.016	1.37
	Indirect	0.027	1.20		*Indirect*	0.141***	3.70
	Total	0.219***	8.51		*Total*	0.158***	3.47
p_3	*Direct*	0.440***	15.68	y_2p_3	*Direct*	0.046	1.59
	Indirect	0.207**	3.26		*Indirect*	−0.025	−0.37
	Total	0.646***	9.75		*Total*	0.022	0.28
p_4	*Direct*	0.054***	4.88	y_2p_4	*Direct*	0.060***	6.67
	Indirect	0.120***	5.84		*Indirect*	0.065**	2.69
	Total	0.174***	8.20		*Total*	0.124***	4.56
y_1^2	*Direct*	0.050***	6.54	y_3p_1	*Direct*	0.024*	2.10
	Indirect	0.118***	4.65		*Indirect*	−0.079*	−2.36
	Total	0.168***	5.85		*Total*	−0.054	−1.42
y_2^2	*Direct*	0.012	1.66	y_3p_3	*Direct*	−0.046*	−1.98
	Indirect	0.087***	3.38		*Indirect*	0.104*	2.01
	Total	0.099***	3.34		*Total*	0.058	0.97
y_3^2	*Direct*	−0.029***	−4.66	y_3p_4	*Direct*	−0.052***	−5.74
	Indirect	−0.045**	−2.76		*Indirect*	0.003	0.11
	Total	−0.074***	−3.75		*Total*	−0.049	−1.66
p_2^2	*Direct*	0.004	0.30	t	*Direct*	−0.012***	−10.75
	Indirect	−0.238***	−5.81		*Indirect*	−0.007***	−9.70
	Total	−0.233***	−4.70		*Total*	−0.018***	−11.55
p_3^2	*Direct*	0.233***	4.24	t^2	*Direct*	0.000	1.67
	Indirect	−0.081	−0.59		*Indirect*	0.000	1.63
	Total	0.152	1.08		*Total*	0.000	1.66
p_4^2	*Direct*	−0.027***	−4.97	y_1t	*Direct*	−0.001	−1.26
	Indirect	−0.084***	−5.62		*Indirect*	0.004**	3.03
	Total	−0.111***	−6.73		*Total*	0.003*	2.30
p_2p_3	*Direct*	0.006	0.14	y_2t	*Direct*	−0.003***	−4.10
	Indirect	0.279*	2.20		*Indirect*	−0.007***	−4.63
	Total	0.285	1.95		*Total*	−0.010***	−5.62
$\ln p_2 \ln p_4$	*Direct*	0.042***	3.87	y_3t	*Direct*	0.004***	6.26
	Indirect	−0.050	−1.77		*Indirect*	0.007***	4.69
	Total	−0.008	−0.24		*Total*	0.012***	6.28
p_3p_4	*Direct*	−0.097***	−4.44	p_2t	*Direct*	0.000	0.04
	Indirect	−0.023	−0.39		*Indirect*	0.012***	3.81
	Total	−0.120*	−2.03		*Total*	0.012**	3.17
y_1y_2	*Direct*	−0.114***	−9.12	p_3t	*Direct*	0.014***	4.88
	Indirect	−0.287***	−6.57		*Indirect*	0.005	0.71
	Total	−0.401***	−8.18		*Total*	0.019*	2.49
y_1y_3	*Direct*	−0.002	−0.22	p_4t	*Direct*	0.000	0.07
	Indirect	−0.052	−1.81		*Indirect*	0.001	0.38
	Total	−0.054	−1.59		*Total*	0.001	0.39
y_2y_3	*Direct*	0.051***	5.01				
	Indirect	0.113***	4.22				
	Total	0.164***	5.26				

SDF denotes the spatial Durbin stochastic frontier model.
*, **, *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

Table 8: Marginal effects from a spatial autoregressive stochastic cost frontier model for purposes of comparison

SAR W_{Cont}^{Row}							
		Marginal Eff.	t-stat			Marginal Eff.	t-stat
y_1	*Direct*	0.465***	6.42	y_1p_2	*Direct*	−0.033***	−3.33
	Indirect	0.218***	5.83		*Indirect*	−0.016**	−3.25
	Total	0.683***	6.36		*Total*	−0.049***	−3.33
y_2	*Direct*	0.937***	8.07	y_1p_3	*Direct*	0.019	0.89
	Indirect	0.437***	8.42		*Indirect*	0.009	0.88
	Total	1.374***	8.49		*Total*	0.028	0.89
y_3	*Direct*	0.037	0.52	y_1p_4	*Direct*	0.014*	2.01
	Indirect	0.017	0.51		*Indirect*	0.006*	1.98
	Total	0.054	0.51		*Total*	0.020*	2.00
p_2	*Direct*	0.206***	28.16	y_2p_2	*Direct*	−0.001	−0.09
	Indirect	0.096***	12.09		*Indirect*	−0.001	−0.09
	Total	0.302***	22.52		*Total*	−0.002	−0.09
p_3	*Direct*	0.468***	17.31	y_2p_3	*Direct*	0.075**	2.63
	Indirect	0.219***	11.93		*Indirect*	0.035*	2.50
	Total	0.687***	17.03		*Total*	0.111**	2.60
p_4	*Direct*	0.055***	6.42	y_2p_4	*Direct*	0.049***	5.26
	Indirect	0.026***	5.37		*Indirect*	0.023***	5.22
	Total	0.081***	6.14		*Total*	0.072***	5.33
y_1^2	*Direct*	0.035***	4.43	y_3p_1	*Direct*	0.027*	2.25
	Indirect	0.016***	4.13		*Indirect*	0.013*	2.20
	Total	0.051***	4.37		*Total*	0.039*	2.24
y_2^2	*Direct*	0.011	1.53	y_3p_3	*Direct*	−0.061**	−2.63
	Indirect	0.005	1.50		*Indirect*	−0.029*	−2.50
	Total	0.016	1.52		*Total*	−0.090**	−2.60
y_3^2	*Direct*	−0.034***	−5.34	y_3p_4	*Direct*	−0.054***	−5.87
	Indirect	−0.016***	−5.05		*Indirect*	−0.025***	−5.58
	Total	−0.050***	−5.32		*Total*	−0.079***	−5.87
p_2^2	*Direct*	0.038**	2.59	t	*Direct*	−0.010***	−14.34
	Indirect	0.018*	2.49		*Indirect*	−0.005***	−9.65
	Total	0.056*	2.57		*Total*	−0.015***	−13.28
p_3^2	*Direct*	0.315***	5.54	t^2	*Direct*	0.000	−0.58
	Indirect	0.147***	5.20		*Indirect*	0.000	−0.58
	Total	0.462***	5.52		*Total*	0.000	−0.58
p_4^2	*Direct*	−0.032***	−5.98	y_1t	*Direct*	−0.001	−1.47
	Indirect	−0.015***	−6.03		*Indirect*	0.000	−1.45
	Total	−0.047***	−6.12		*Total*	−0.001	−1.46
p_2p_3	*Direct*	−0.043	−0.97	y_2t	*Direct*	−0.001	−1.17
	Indirect	−0.021	−0.98		*Indirect*	0.000	−1.15
	Total	−0.064	−0.97		*Total*	−0.001	−1.16
p_2p_4	*Direct*	0.055***	4.72	y_3t	*Direct*	0.003***	4.67
	Indirect	0.026***	4.33		*Indirect*	0.001***	4.88
	Total	0.080***	4.64		*Total*	0.004***	4.79
p_3p_4	*Direct*	−0.134***	−6.24	p_2t	*Direct*	−0.001	−0.60
	Indirect	−0.063***	−5.75		*Indirect*	0.000	−0.61
	Total	−0.197***	−6.20		*Total*	−0.001	−0.60
y_1y_2	*Direct*	−0.104***	−7.99	p_3t	*Direct*	0.016***	6.05
	Indirect	−0.049***	−6.82		*Indirect*	0.008***	5.61
	Total	−0.152***	−7.80		*Total*	0.024***	6.02
y_1y_3	*Direct*	0.024*	2.55	p_4t	*Direct*	−0.002**	−2.79
	Indirect	0.011*	2.51		*Indirect*	−0.001**	−2.64
	Total	0.035*	2.55		*Total*	−0.003**	−2.75
y_2y_3	*Direct*	0.035**	3.22				
	Indirect	0.017**	3.14				
	Total	0.052**	3.22				

SAR denotes the spatial autoregressive stochastic frontier model.
*, **, *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

Table 9: Average net time-invariant, net time-variant and gross time-variant cost efficiency scores and rankings

Panel A: NSF						
State	Av NIE Score	Av NIE Rank	Av NVE Score	Av NVE Rank	Av GVE Score	Av GVE Rank
California	0.999	45	0.950	33	0.950	33
Iowa	0.999	47	0.953	13	0.952	13
Texas	0.999	48	0.954	8	0.953	8
Oklahoma	0.999	33	0.952	22	0.952	22
Kentucky	0.999	29	0.952	24	0.952	24
Mississippi	0.999	20	0.952	27	0.951	27
Idaho	0.999	15	0.952	25	0.952	25
Nevada	0.999	2	0.949	43	0.948	43
New Hampshire	0.999	4	0.950	37	0.949	37
Rhode Island	0.999	1	0.947	45	0.947	45
Sample	0.999		0.951		0.951	
Panel B: SAR W^{Row}_{Cont}						
State	Av NIE Score	Av NIE Rank	Av NVE Score	Av NVE Rank	Av GVE Score	Av GVE Rank
California	0.787	48	0.674	21	0.530	33
Iowa	0.842	44	0.637	28	0.536	31
Texas	0.799	47	0.637	29	0.509	41
Oklahoma	0.873	27	0.626	34	0.547	29
Kentucky	0.876	23	0.669	23	0.586	22
Mississippi	0.877	21	0.595	38	0.522	37
Idaho	0.866	34	0.695	18	0.602	19
Nevada	0.923	1	0.627	32	0.579	23
New Hampshire	0.898	9	0.709	16	0.637	14
Rhode Island	0.919	2	0.840	2	0.772	1
Sample	0.873		0.666		0.582	
Panel C: SDF W^{Row}_{Cont}						
State	Av NIE Score	Av NIE Rank	Av NVE Score	Av NVE Rank	Av GVE Score	Av GVE Rank
California	1.000	47	0.674	21	0.674	21
Iowa	1.000	46	0.637	28	0.637	28
Texas	1.000	47	0.637	29	0.637	29
Oklahoma	1.000	22	0.626	34	0.626	34
Kentucky	1.000	22	0.670	22	0.670	22
Mississippi	1.000	15	0.595	38	0.595	38
Idaho	1.000	22	0.695	18	0.694	18
Nevada	1.000	2	0.627	32	0.627	32
New Hampshire	1.000	4	0.710	16	0.709	16
Rhode Island	1.000	1	0.839	2	0.839	2
Sample	1.000		0.666		0.666	

NSF denotes the non-spatial stochastic frontier model, SAR denotes the spatial autoregressive stochastic frontier model and SDF denotes the spatial Durbin stochastic frontier model.

NIE, NVE and GVE denote net time-invariant, net time-variant and gross time-variant cost efficiencies.

Table 10: Average absolute direct, indirect and total gross time-variant cost efficiency scores and rankings

Panel A: W_{Cont}^{Row} SAR

State	Av GVE^{Dir} Score	Av GVE^{Dir} Rank	Av GVE_{To}^{Ind} Score	Av GVE_{To}^{Ind} Rank	Av GVE_{From}^{Ind} Score	Av GVE_{From}^{Ind} Rank	Av GVE_{To}^{Tot} Score	Av GVE_{To}^{Tot} Rank	Av GVE_{From}^{Tot} Score	Av GVE_{From}^{Tot} Rank
California	0.55	33	0.30	7	0.21	37	0.84	30	0.73	43
Iowa	0.55	31	0.28	25	0.33	14	0.83	31	0.86	25
Texas	0.52	40	0.27	36	0.24	33	0.79	40	0.74	5
Oklahoma	0.56	29	0.26	42	0.32	18	0.82	32	0.88	11
Kentucky	0.60	22	0.30	5	0.40	8	0.90	19	0.99	39
Mississippi	0.54	37	0.27	35	0.24	34	0.80	37	0.76	16
Idaho	0.62	19	0.29	10	0.46	2	0.92	17	1.06	12
Nevada	0.60	23	0.28	17	0.34	11	0.88	23	0.94	4
New Hampshire	0.68	12	0.25	44	0.34	12	0.93	15	1.08	23
Rhode Island	0.80	1	0.31	2	0.16	43	1.10	1	1.00	41
Sample	0.60		0.28		0.28		0.88		0.88	

Panel B: W_{Cont}^{Row} SDF

State	Av GVE^{Dir} Score	Av GVE^{Dir} Rank	Av GVE_{To}^{Ind} Score	Av GVE_{To}^{Ind} Rank	Av GVE_{From}^{Ind} Score	Av GVE_{From}^{Ind} Rank	Av GVE_{To}^{Tot} Score	Av GVE_{To}^{Tot} Rank	Av GVE_{From}^{Tot} Score	Av GVE_{From}^{Tot} Rank
California	0.70	21	0.41	10	0.28	39	1.11	20	0.98	37
Iowa	0.66	30	0.40	19	0.46	13	1.06	27	1.12	18
Texas	0.66	28	0.37	39	0.33	32	1.03	34	1.00	35
Oklahoma	0.65	34	0.37	38	0.45	14	1.02	35	1.10	21
Kentucky	0.69	23	0.42	5	0.55	6	1.12	17	1.24	9
Mississippi	0.62	38	0.37	36	0.31	33	0.99	38	0.93	40
Idaho	0.73	18	0.41	13	0.62	3	1.14	13	1.35	4
Nevada	0.66	31	0.40	17	0.45	17	1.06	28	1.11	19
New Hampshire	0.77	13	0.35	45	0.53	7	1.13	16	1.31	7
Rhode Island	0.88	2	0.42	3	0.26	42	1.30	2	1.14	16
Sample	0.69		0.39		0.39		1.09		1.09	

NSF denotes the non-spatial stochastic frontier model, SAR denotes the spatial autoregressive stochastic frontier model and SDF denotes the spatial Durbin stochastic frontier model.
Other notation is as specified in the text.

Table 11: Average partitioned absolute direct, indirect and total gross time-variant cost efficiency scores and rankings

		W_{Cont}^{Row} SDF									
State	Order	Av GVE^{Dir} Score	Av GVE^{Dir} Rank	Av GVE_{To}^{Ind} Score	Av GVE_{To}^{Ind} Rank	Av GVE_{From}^{Ind} Score	Av GVE_{From}^{Ind} Rank	Av GVE_{To}^{Tot} Score	Av GVE_{To}^{Tot} Rank	Av GVE_{From}^{Tot} Score	Av GVE_{From}^{Tot} Rank
California	W^0	0.67	21	0.00	1	0.00	1	0.67	21	0.67	21
	W^1	0.00	1	0.27	9	0.18	39	0.27	9	0.18	39
	W^2	0.02	20	0.08	26	0.06	37	0.10	23	0.08	36
	W^3	0.00	21	0.04	9	0.03	41	0.04	9	0.03	39
	W^4	0.00	16	0.01	23	0.01	40	0.02	21	0.01	37
Iowa	W^0	0.64	28	0.00	1	0.00	1	0.64	28	0.64	28
	W^1	0.00	1	0.26	24	0.29	16	0.26	24	0.29	16
	W^2	0.02	37	0.08	2	0.10	10	0.10	16	0.12	11
	W^3	0.00	16	0.04	18	0.04	14	0.04	20	0.05	14
	W^4	0.00	37	0.01	4	0.02	11	0.02	17	0.02	13
Kentucky	W^0	0.67	22	0.00	1	0.00	1	0.67	22	0.67	22
	W^1	0.00	1	0.28	5	0.34	9	0.28	5	0.34	9
	W^2	0.02	36	0.08	8	0.12	4	0.10	21	0.14	5
	W^3	0.00	18	0.04	20	0.05	4	0.04	21	0.06	5
	W^4	0.00	34	0.01	12	0.02	4	0.02	27	0.02	4
Mississippi	W^0	0.59	38	0.00	1	0.00	1	0.59	38	0.59	38
	W^1	0.00	1	0.24	38	0.20	35	0.24	38	0.20	35
	W^2	0.02	32	0.08	32	0.06	30	0.09	42	0.08	35
	W^3	0.00	42	0.03	35	0.03	32	0.04	44	0.03	33
	W^4	0.00	36	0.01	35	0.01	32	0.01	45	0.01	34
Nevada	W^0	0.63	32	0.00	1	0.00	1	0.63	32	0.63	32
	W^1	0.00	1	0.27	16	0.29	15	0.27	16	0.29	15
	W^2	0.02	25	0.08	21	0.09	13	0.10	18	0.12	12
	W^3	0.00	3	0.04	33	0.04	21	0.04	17	0.04	19
	W^4	0.00	11	0.01	27	0.01	19	0.02	18	0.02	17
New Hampshire	W^0	0.71	16	0.00	1	0.00	1	0.71	16	0.71	16
	W^1	0.00	1	0.25	35	0.42	3	0.25	35	0.42	3
	W^2	0.05	1	0.05	48	0.04	43	0.10	13	0.10	23
	W^3	0.00	43	0.04	13	0.05	6	0.04	25	0.05	7
	W^4	0.00	1	0.01	48	0.01	35	0.02	5	0.01	22
Sample	W^0	0.67		0.00		0.00		0.67		0.67	
	W^1	0.00		0.26		0.26		0.26		0.26	
	W^2	0.02		0.08		0.08		0.10		0.10	
	W^3	0.00		0.04		0.04		0.04		0.04	
	W^4	0.00		0.01		0.01		0.02		0.02	

NSF denotes the non-spatial stochastic frontier model, SAR denotes the spatial autoregressive stochastic frontier model and SDF denotes the spatial Durbin stochastic frontier model.
Other notation is as specified in the text.