Generalized Moments Estimation for Spatial Panel Data: Indonesian Rice Farming

Viliam Druska

William C. Horrace*

March, 2003

Abstract:
We consider estimation of a panel data model where disturbances are *spatially correlated* in the cross-sectional dimension, based on geographic or economic proximity. When the time dimension of the data is large, spatial correlation parameters may be consistently estimated. When the time dimension is small (the usual panel data case), we develop an estimator that extends the cross-sectional model of Kelejian and Prucha. This approach is applied in a stochastic frontier framework to a panel of Indonesian rice farms where spatial correlations represent productivity shock spillovers, based on geographic proximity and weather. These spillovers affect farm-level efficiency estimation and ranking.

Keywords: autocorrelation, productivity, stochastic frontiers, economic spillovers.
JEL Codes: C21, C23, D24

Viliam Druska is a Ph.D. candidate at the Center for Economic Research and Graduate Education, Charles University, Prague. William C. Horrace is Associate Professor at the Center for Policy Research and Department of Economics, Syracuse University, and Research Economist at the National Bureau of Economic Research, Cambridge, MA.

* Corresponding author. Center for Policy Research, Syracuse University, 426 Eggers Hall, Syracuse, NY 13244-1020, USA. Voice: (315) 433-9061, Fax: (315) 443-1081. Email: whorrace@maxwell.syr.edu. The second author gratefully acknowledges the support of the National Science Foundation, the National Bureau of Economic Research, and the Office of the Vice Chancellor of Syracuse University. Professor Jeffrey Racine provided the density estimates. Martha Bonney and Toni Metzger provided editorial assistance. An anonymous referee provided excellent comments.

Generalized Moments Estimation for Spatial Panel Data: Indonesian Rice Farming

Introduction

Much has been written about spatial dependence in cross-sectional economic data that can be distinguished by absolute or relative location. For example, data on employment or wealth can be organized by county, state, census tract, or country, and spatial dependence can be modeled across these units. Anselin provides an excellent textbook treatment of the analysis of spatially dependent data. Theoretical or empirical spatial issues have also been addressed in Case; Conley; Delong and Summers; Dubin; Fishback, Horrace and Kantor; Kelejian and Robinson; Moulton; Quah; and Topa. These cross-sectional specifications address the important phenomena of spatial aggregation, infrastructure effects, and economic spillovers, to name a few.

Kelejian and Prucha consider generalized moments estimation of regression models which allow spatial autocorrelation of *disturbances* across cross-sectional units. Estimation hinges on the ex ante specification of a "spatial weighting matrix" in the regression error. The form of the weighting matrix is at the discretion of the analyst, but often it can be based on some underlying economic, geographic, or meteorological theory. "Of course, if panel data are available one can consider, e.g., a seemingly unrelated regression model, or an error component model to permit for cross-sectional correlation, and estimate the cross-sectional correlations via the time dimension of the panel if the time dimension is large" (Kelejian and Prucha, footnote 2, p. 509). Unfortunately, in the usual panel data case, the time dimension is *small* (fixed), so consistent estimation of the cross-sectional correlations is typically not justified.

This article extends the Kelejian and Prucha estimator to the usual panel data case, based on certain restrictions on the evolution of spatial dependence over time. It is important to stress that the panel data theory presented is for the case where T is fixed; consequently, the current discussion also hinges on the ex ante specification of a spatial weighting matrix. Once we allow the time dimension to grow, the specification of the weighting matrix becomes unnecessary, as the estimation techniques presented herein become empirically inferior to approaches that rely on T asymptotics, such as seemingly unrelated regression models or error component models.

We apply these spatial techniques to the stochastic frontier model in which a common production function and farm-level technical efficiencies are estimated for a sample of farm inputs and outputs. Cross-sectional estimation of these models is due to Aigner, Lovell and Schmidt; and Meeusen and van den Broeck, while panel estimation is due to Schmidt and Sickles. Our concern is, of course, the panel specification, and we select a panel of 171 Indonesian rice farms observed over six periods for our example. Output is rice, and inputs are things like seed, fertilizer, and land acreage. The time dimension of the data is small, so consistent estimation of cross-sectional correlations in the error process is not justified. Consequently, we specify a spatial weighting scheme in the error process that allows for spillovers across farms based on geographic proximity and weather conditions. The results indicate that spatial correlations exist in the data and have an impact on the magnitude and variability of the production function and technical efficiency estimates obtained.

A Panel Model with Spatial Disturbances

Consider the standard fixed effect (FE) model

$$y_{it} = \alpha_i + \beta'\mathbf{x}_{it} + u_{it}; \quad i = 1, ..., N; \quad t = 1, ..., T,$$

where β is $(k\times1)$ and $\mathbf{x}_{it}$ is $(1\times k)$. Here we assume that T is fixed, so we cannot rely on T-asymptotics. Forming vectors of observations in i, the model becomes

(1) $\quad \mathbf{y}_t = \alpha + \mathbf{x}_t\beta + \mathbf{u}_t, \quad t = 1, ..., T,$

where $\mathbf{y}_t' = [y_{1t}, ..., y_{Nt}]$, $\alpha' = [\alpha_1, ..., \alpha_N]$, $\mathbf{x}_t' = [\mathbf{x}_{1t}', ..., \mathbf{x}_{Nt}']$, and $\mathbf{u}_t' = [u_{1t}, ..., u_{Nt}]$. Now suppose that the error term is spatially lagged such that

(2) $\quad \mathbf{u}_t = \rho_t\mathbf{M}_t\mathbf{u}_t + \varepsilon_t, \quad t = 1, ..., T,$

where ρ_t is a scalar, spatial autoregressive parameter, $\mathbf{M}_t$ is a $(N\times N)$ spatial weighting matrix of known constants (diagonal elements equal to zero), which captures the spatial correlations across cross-sectional units, and ε_t $(N\times1)$ is a zero-mean disturbance. (Later we allow for a time-invariant spatial parameter and weighting matrix.) Elements of $\mathbf{M}_t$ are m_{ijt}, and they are chosen based on some geographic or economic proximity measure such as contiguity or physical, economic or climatic distances or dissimilarities. For example, in the application, we select m_{ijt} to be the inverse of the physical distance (1/km) between unit i and unit j in time period t.

The application of interest is the stochastic frontier model, where y_{it} and x_{it} are the productive output and exogenous inputs, respectively, of farm i in period t. Stochastic frontier models specify output as a linear function of a) farm-level technical (in)efficiency (an unobserved factor imputed to each farm, embodied in α_i), b) a representative log-production function (deterministic, within the control of each farm, and represented by $\mathbf{x}_t\beta$) and c) productivity shocks (random, out of the farmer's control, and represented by $\mathbf{u}_t$). Therefore, equation (1) is a stochastic frontier specification. When augmented by equation (2), the specification implies that, in each period t, productivity shocks are correlated across i, and specifically that the productive output of farm i is a function of the spatial lag of productivity shocks, $\rho_t \mathbf{M}_t \mathbf{u}_t$, experienced by other farms in the sample. This would seem reasonable if productivity shocks included geographic or climatic unobservables that affected farms in similar ways but were location or climate specific (e.g., unmeasured rainfall, temperature and sunlight). Notice that there is no spatial lag of $\mathbf{y}_t$ on the right-hand side of equation (1). Therefore, the specification *implicitly assumes* that, in each period t, the productive output of farm i is *not* a function of the output of other farms in the sample. This seems reasonable if the production function is viewed as a purely deterministic (engineering) process, where the farmer controls all inputs. We need the additional assumptions:

> *Assumption 1:* The elements of $\boldsymbol{\varepsilon}_t$ are independently and identically distributed with zero-mean and finite variance σ_t^2, the fourth moment of $\boldsymbol{\varepsilon}_t$ is finite, and $\boldsymbol{\varepsilon}_t$ is independent of $\boldsymbol{\varepsilon}_s$, $\forall\ t \neq s$.

> *Assumption 2:* All diagonal elements of $\mathbf{M}_t$ are zero. The matrix $(\mathbf{I}_N - \rho_t \mathbf{M}_t)$ is non-singular. $|\rho_t| < 1$.

Notice that under Assumptions 1 and 2, $\mathbf{u}_t = (\mathbf{I}_N - \rho_t \mathbf{M}_t)^{-1} \boldsymbol{\varepsilon}_t$, so $E(\mathbf{u}_t) = 0$ for all t, but $E(\mathbf{u}_t \mathbf{u}_t')$ has a general, non-spherical structure, which is a function of ρ_t, $\mathbf{M}_t$, and σ_t^2. Since $\mathbf{M}_t$ is known, $E(\mathbf{u}_t \mathbf{u}_t')$ is known up to ρ_t and σ_t^2, parameters which we will ultimately estimate. Estimation of ρ_t and σ_t^2 allows feasible and efficient estimation of equation (1). Also, notice that if $\rho_t = \rho$, $\mathbf{M}_t = \mathbf{M}$, and $\sigma_t^2 = \sigma^2$, then $E(\mathbf{u}_t \mathbf{u}_t')$ is a constant, which can be consistently estimated as $T \rightarrow \infty$. Here, we assume that T is fixed, so consistent estimation of $E(\mathbf{u}_t \mathbf{u}_t')$ is unreasonable, and we must assume that $\mathbf{M}_t$ is known to identify an estimate of equation (1). For now, assume that ρ_t and σ_t^2 are known. Forming vectors in t from the vectors of observations in i,

(3) $\mathbf{y} = \iota_T \otimes \alpha + \mathbf{x}\beta + \mathbf{u}, \quad \mathbf{u} = (\rho^* \otimes \mathbf{I}_N)\, \mathbf{M}^* \boldsymbol{u} + \varepsilon,$

where ι_T is a T dimensional column vector of ones, $\mathbf{y}' = [\mathbf{y}_1', ..., \mathbf{y}_T']$, $\mathbf{x}' = [\mathbf{x}_1', ..., \mathbf{x}_T']$, $\mathbf{u}' = [\mathbf{u}_1', ..., \mathbf{u}_T']$, and

$$\mathbf{M}^* = \begin{bmatrix} \mathbf{M}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{M}_T \end{bmatrix}, \qquad \rho^* = \begin{bmatrix} \rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \rho_T \end{bmatrix}.$$

Notice that

$$\mathrm{E}(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_1^2 \mathbf{I}_N & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_T^2 \mathbf{I}_N \end{bmatrix},$$

so the disturbance in equation (2) is heteroskedastic. Define $\mathbf{\Phi}_t = (\mathbf{I}_N - \rho_t \mathbf{M}_t)/\sigma_t$; then we can pre-multiply the model in equations (1) and (2) to get

(4) $\mathbf{y}_t^* = \boldsymbol{\alpha}_t^* + \mathbf{x}_t^* \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t^*,$

where $\mathbf{y}_t^* = \mathbf{\Phi}_t \mathbf{y}_t$, $\mathbf{x}_t^* = \mathbf{\Phi}_t \mathbf{x}_t$, $\boldsymbol{\alpha}_t^* = \mathbf{\Phi}_t \boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}_t^* = \mathbf{\Phi}_t \mathbf{u}_t = \boldsymbol{\varepsilon}_t / \sigma_t$. Stacking observation in t,

(5) $\mathbf{y}^* = \alpha^* + \mathbf{x}^* \beta + \varepsilon^*,$

where $\boldsymbol{\alpha}^{*\prime} = [\boldsymbol{\alpha}_1^{*\prime}, ..., \boldsymbol{\alpha}_T^{*\prime}]$, a TN dimensional vector. Equation (5) possess a "well-behaved" disturbance; that is, $\mathrm{E}(\varepsilon^*) = 0$ and $\mathrm{E}(\varepsilon^* \varepsilon^{*\prime}) = \mathbf{I}_{TN}$. Identification of any estimates of the parameters in equation (5) hinges on estimation of the unknown parameters $\mathbf{M}_t$, ρ_t, and σ_t^2, which will be undertaken later. The Kelejian and Prucha cross-sectional procedure could be directly applied to equation (4) T times over N observations to recover estimates of ρ_t and σ_t^2 for known $\mathbf{M}_t$. These estimates could then be used to estimate the parameters in equation (5).[1] We refer to this estimation technique as "unrestricted estimation." Our application implies some equality restrictions on the model in equation (5). In particular, our definitions of spatial dependence are based on distinct physical characteristics of the farming villages on the island of Java (longitude, latitude, infrastructure, etc.), which are certainly constant over the short time period of the data (six years). Therefore, we impose some equality restrictions on equation (5) to identify alternative estimators of the models parameters.

A Fully Restricted Specification

One obvious equality restriction is to assume that some subsets of the weighting matrices, autoregressive parameters and variance parameters are equal. As an extreme case we could assume that $\mathbf{M}_1 = \ldots = \mathbf{M}_T = \mathbf{M}$, $\rho_1 = \ldots = \rho_T = \rho$, and $\sigma_1^2 = \ldots = \sigma_T^2 = \sigma^2$, implying $\mathbf{\Phi}_1 = \ldots = \mathbf{\Phi}_T = \mathbf{\Phi}$. Then $\boldsymbol{\alpha}_t^* = \Phi\boldsymbol{\alpha}^*$ in equation (4), and $\alpha^* = \iota_T \otimes \mathbf{\Phi}\alpha$ in equation (5). Of course, the error term ε of equation (3) is no longer heteroskedastic; it has variance matrix $E(\varepsilon\varepsilon') = \sigma^2 \mathbf{I}_{TN}$, so $\mathbf{\Phi}$ need not be a function of σ for efficiency. Fixed effect estimation of equation (5) under this full restriction will then be efficient for α^* and β, if ρ and σ^2 are known, and if the equality restriction is true. It is also consistent for fixed T as $N \rightarrow \infty$. Additionally, an estimate of α can be recovered by transforming the estimate of α^* with $\mathbf{\Phi}$. Of course, ρ and σ^2 are not known, so the challenge is to consistently estimate them, so that equation (5) can be *feasibly* estimated; this is undertaken in the next section.

A Partially Restricted Specification

As another example of a reasonable restriction on the parameters of the model, briefly consider the empirical example. We observe $N = 171$ Indonesian rice farms over $T = 6$ periods. Periods 1, 3 and 5 are "wet or rainy seasons" and periods 2, 4 and 6 are "dry seasons". It may be reasonable to suspect that $\rho_1 = \rho_3 = \rho_5 = \rho_W$ (wet) and $\rho_2 = \rho_4 = \rho_6 = \rho_D$ (dry), and similarly for $\mathbf{M}_t$, σ_t^2, and $\mathbf{\Phi}_t$. (This may be true on the island of Java, since during the rainy season many roads in the low country are impassable, and hence spillovers based on infrastructure are potentially diminished.) Then

$$\alpha^* = [\, (\Phi_W\, \alpha)' \, (\Phi_D\alpha)' \, (\Phi_W\, \alpha)' \, (\Phi_D\alpha)' \, (\Phi_W\, \alpha)' \, (\Phi_D\alpha)']' \,,$$

in equation (5), a TN dimensional column vector consisting of $2N$ parameters. The system in (5) then consists of $2N + k$ parameters and can effectively be treated as $2\times171 = 342$ farms observed over $6/2 = 3$ periods, so fixed effect estimation of equation (5) is feasible, since it has been assumed that realizations of the error ε_t are independent across both t and i. Of course, there will be an efficiency loss in the estimate of α^* relative to the fully restricted estimate, since the time series dimension has been effectively cut in half from 6 to 3, but the slope parameter β will still be efficient (and consistent in N) since it is still based on the same number of observations, TN. Again the challenge is estimation of ρ_W, ρ_D, σ_W^2, and σ_D^2, which is undertaken in the following section.

Feasible Estimation

Kelejian and Prucha develop a moments estimator of the parameters ρ_t and σ_t^2 in the cross-sectional setting ($T = 1$). We now generalize their results for the case where ρ_t and σ_t^2 are different across t.[2] Using their notation, let $\tilde{\mathbf{u}}_t$ be a predictor of $\mathbf{u}_t$ from the fixed effect (or within) regression implied by equation (1), ignoring equation (2). That is, $\tilde{\mathbf{u}}_t$ converges in distribution to the random variable $\mathbf{u}_t$. Additionally, let $\tilde{\bar{\mathbf{u}}}_t = \mathbf{M}_t\tilde{\mathbf{u}}_t$, $\tilde{\bar{\bar{\mathbf{u}}}}_t = \mathbf{M}_t\tilde{\bar{\mathbf{u}}}_t$, $\bar{\boldsymbol{\varepsilon}}_t = \mathbf{M}_t\boldsymbol{\varepsilon}_t$, and $\bar{\bar{\boldsymbol{\varepsilon}}}_t = \mathbf{M}_t\bar{\boldsymbol{\varepsilon}}_t$. Consider the following $3T$ moment conditions implied by equations (1) and (2) and Assumptions 1 and 2,

$$E[N^{-1}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t] = \sigma_t^2, \qquad E[N^{-1}\bar{\boldsymbol{\varepsilon}}_t'\bar{\boldsymbol{\varepsilon}}_t] = \sigma_t^2 N^{-1} tr(\mathbf{M}_t'\mathbf{M}_t), \qquad E[N^{-1}\bar{\boldsymbol{\varepsilon}}_t'\boldsymbol{\varepsilon}_t] = 0,$$

$t = 1, \ldots, T$. Noting that $\boldsymbol{\varepsilon}_t = (\mathbf{I}_N - \rho_t\mathbf{M}_t)\mathbf{u}_t$, these moment conditions imply the following system of $3T$ equations,

$$\boldsymbol{\Gamma}_t[\rho_t \quad \rho_t^2 \quad \sigma_t^2]' - \boldsymbol{\gamma}_t = \mathbf{0},$$

where,

$$\boldsymbol{\Gamma}_t = \begin{bmatrix} \frac{2}{N}E(\mathbf{u}_t'\bar{\mathbf{u}}_t) & \frac{-1}{N}E(\bar{\mathbf{u}}_t'\bar{\mathbf{u}}_t) & 1 \\ \frac{2}{N}E(\bar{\bar{\mathbf{u}}}_t'\bar{\mathbf{u}}_t) & \frac{-1}{N}E(\bar{\bar{\mathbf{u}}}_t'\bar{\bar{\mathbf{u}}}_t) & \frac{1}{N}tr(\mathbf{M}_t'\mathbf{M}_t) \\ \frac{1}{N}E(\mathbf{u}_t'\bar{\bar{\mathbf{u}}}_t + \bar{\mathbf{u}}_t'\bar{\mathbf{u}}_t) & \frac{-1}{N}E(\bar{\mathbf{u}}_t'\bar{\bar{\mathbf{u}}}_t) & 0 \end{bmatrix}, \qquad \gamma_t = \begin{bmatrix} \frac{1}{N}E(\mathbf{u}_t'\mathbf{u}_t) \\ \frac{1}{N}E(\bar{\mathbf{u}}_t'\bar{\mathbf{u}}_t) \\ \frac{1}{N}E(\mathbf{u}_t'\bar{\mathbf{u}}_t) \end{bmatrix},$$

$t = 1, \ldots, T$. The sample analogs based on $\tilde{\mathbf{u}}_t$ are

$$\text{(6)} \qquad \mathbf{G}_t[\rho_t \quad \rho_t^2 \quad \sigma_t^2]' - \mathbf{g}_t = \mathbf{v}_t(\rho_t, \sigma_t^2),$$

$$\mathbf{G}_t = \begin{bmatrix} \frac{2}{N}\tilde{\mathbf{u}}_t'\tilde{\bar{\mathbf{u}}}_t & \frac{-1}{N}\tilde{\bar{\mathbf{u}}}_t'\tilde{\bar{\mathbf{u}}}_t & 1 \\ \frac{2}{N}\tilde{\bar{\bar{\mathbf{u}}}}_t'\tilde{\bar{\mathbf{u}}}_t & \frac{-1}{N}\tilde{\bar{\bar{\mathbf{u}}}}_t'\tilde{\bar{\bar{\mathbf{u}}}}_t & \frac{1}{N}tr(\mathbf{M}_t'\mathbf{M}_t) \\ \frac{1}{N}\tilde{\mathbf{u}}_t'\tilde{\bar{\bar{\mathbf{u}}}}_t + \tilde{\bar{\mathbf{u}}}_t'\tilde{\bar{\mathbf{u}}}_t & \frac{-1}{N}\tilde{\bar{\mathbf{u}}}_t'\tilde{\bar{\bar{\mathbf{u}}}}_t & 0 \end{bmatrix}, \qquad \mathbf{g}_t = \begin{bmatrix} \frac{1}{N}\tilde{\mathbf{u}}_t'\tilde{\mathbf{u}}_t \\ \frac{1}{N}\tilde{\bar{\mathbf{u}}}_t'\tilde{\bar{\mathbf{u}}}_t \\ \frac{1}{N}\tilde{\mathbf{u}}_t'\tilde{\bar{\mathbf{u}}}_t \end{bmatrix},$$

$t = 1, \ldots, T$. Here $\mathbf{v}_t$ is the usual error associated with a sample of statistical realizations (i.e., each element will ultimately be squared, summed, then minimized by selecting parameters optimally). The system consists of $3T$ equations and $3T$ unknowns, but the system is actually T separate subsystems of 3 equations and 3 unknowns. If these T subsystems satisfy Assumptions 1 and 2 above and Assumptions 3, 4. and 5 of Kelejian and Prucha, then Theorem 1 of Kelejian and Prucha is applicable to the individual subsystems.[3] That is, $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ which solve the nonlinear optimization

(7) $$(\hat{\rho}_t, \hat{\sigma}_t^2) = \arg\min_{r,s^2}[\mathbf{v}_t(r,s^2)'\mathbf{v}_t(r,s^2) : s^2 \geq 0],$$

are consistent for ρ_t and σ_t^2 as $N \to \infty$. For a proof see Kelejian and Prucha. Let

$$\hat{\mathbf{\Phi}}_t = (\mathbf{I}_N - \hat{\rho}_t \mathbf{M}_t) / \sqrt{\hat{\sigma}_t^2}\ .$$

(We could substitute $\hat{\mathbf{\Phi}}_t$ for $\mathbf{\Phi}_t$ and estimate equation (5), but we ultimately chose to restrict the model.) Let us call the $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ *unrestricted* estimates.

Feasible Estimation of the Fully Restricted System

If we can assume that $\mathbf{M}_1 = \ldots = \mathbf{M}_T = \mathbf{M}$, $\rho_1 = \ldots = \rho_T = \rho$, and $\sigma_1^2 = \ldots = \sigma_T^2 = \sigma^2$ as before, then we can impose the assumptions in equation (6) and estimate $\hat{\rho}_t$ and $\hat{\sigma}_t^2$, $t = 1, \ldots, T$ as above.[4] Then average estimates of ρ and σ^2 are

(8) $$\hat{\rho} = T^{-1}\sum_t \hat{\rho}_t \text{ and } \hat{\sigma}^2 = T^{-1}\sum_t \hat{\sigma}_t^2\ .$$

We shall call these estimates the *fully restricted average* estimates. The estimates will be consistent as $N \to \infty$, as long as the restriction is true. These are two-stage estimates, where in the first stage unrestricted estimates are calculated ($\hat{\rho}_t$ and $\hat{\sigma}_t^2$, $t = 1, \ldots, T$), and the restriction is imposed in the second stage of averaging over t. Since the estimates are based on the unrestricted estimates, they do not exploit all the information in the data set simultaneously. That is, each $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ is calculated from one of T separate subsamples of the data. These estimates imply that

$$\hat{\mathbf{\Phi}} = (\mathbf{I}_N - \hat{\rho}\mathbf{M}),$$

which can be substituted into equation (5). Then fixed effect estimation of equation (5) with $\boldsymbol{\alpha}^* = \boldsymbol{\iota}_T \otimes \hat{\boldsymbol{\Phi}}\boldsymbol{\alpha}$ is consistent for $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}$.

If we can a) impose the restriction, b) estimate the parameters in a single step, and c) do so such that the data is not divided into T subsamples, then the resulting parameter estimates should be more efficient than the average fully restricted estimates. One such estimate is based on the moment conditions:

$$E[(TN)^{-1}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}] = \sigma^2, \quad E[(TN)^{-1}\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}] = \sigma^2 N^{-1} tr(\mathbf{M}'\mathbf{M}),\ E[(TN)^{-1}\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\varepsilon}] = 0,$$

where $\bar{\boldsymbol{\varepsilon}} = \mathbf{M}^*\boldsymbol{\varepsilon}$, and $\bar{\bar{\boldsymbol{\varepsilon}}} = \mathbf{M}^*\bar{\boldsymbol{\varepsilon}}$.[5] Let $\tilde{\mathbf{u}}$ be a predictor of $\mathbf{u}$ from the fixed effect (or within) regression implied by equation (3), $\bar{\tilde{\mathbf{u}}} = \mathbf{M}^*\tilde{\mathbf{u}}$ and $\bar{\bar{\tilde{\mathbf{u}}}} = \mathbf{M}^*\bar{\tilde{\mathbf{u}}}$, equation (6) becomes

$$\text{(9)} \qquad \mathbf{G}[\rho \quad \rho^2 \quad \sigma^2]' - \mathbf{g} = \mathbf{v}(\rho, \sigma^2),$$

where,

$$\mathbf{G} = \begin{bmatrix} \frac{2}{TN}\tilde{\mathbf{u}}'\bar{\tilde{\mathbf{u}}} & \frac{-1}{TN}\bar{\tilde{\mathbf{u}}}'\bar{\tilde{\mathbf{u}}} & 1 \\ \frac{2}{TN}\bar{\tilde{\mathbf{u}}}'\bar{\bar{\tilde{\mathbf{u}}}} & \frac{-1}{TN}\bar{\bar{\tilde{\mathbf{u}}}}'\bar{\bar{\tilde{\mathbf{u}}}} & \frac{1}{N}tr(\mathbf{M}'\mathbf{M}) \\ \frac{1}{TN}\tilde{\mathbf{u}}'\bar{\bar{\tilde{\mathbf{u}}}} + \bar{\tilde{\mathbf{u}}}'\bar{\tilde{\mathbf{u}}} & \frac{-1}{TN}\bar{\tilde{\mathbf{u}}}'\bar{\bar{\tilde{\mathbf{u}}}} & 0 \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \frac{1}{TN}\tilde{\mathbf{u}}'\tilde{\mathbf{u}} \\ \frac{1}{TN}\bar{\tilde{\mathbf{u}}}'\bar{\tilde{\mathbf{u}}} \\ \frac{1}{TN}\tilde{\mathbf{u}}'\bar{\tilde{\mathbf{u}}} \end{bmatrix}.$$

The system consists of 3 equations and 3 unknowns and is exactly the Kelejian and Prucha result. Then estimates $\tilde{\rho}$ and $\tilde{\sigma}^2$ follow from

$$\text{(10)} \qquad (\tilde{\rho}, \tilde{\sigma}^2) = \arg\min_{r,s^2}[\mathbf{v}(r,s^2)'\mathbf{v}(r,s^2) : s^2 \geq 0].$$

We shall call these the *fully restricted moment estimates* (to differentiate them from the fully restricted average estimates). The potential efficiency gain over the estimates $\hat{\rho}$ and $\hat{\sigma}^2$ hinges on the fact that equation (9) exploits the information contained in TN observations and imposes a hypothetically correct restriction, while equation (6) exploits that contained in only N observations over $t = 1, \ldots, T$, and no restriction. Again, $\tilde{\rho}$ and $\tilde{\sigma}^2$ imply $\tilde{\mathbf{\Phi}}$, which can be inserted into equation (5); then fixed effect estimation of equation (5) with $\boldsymbol{\alpha}^* = \iota_T \otimes \tilde{\mathbf{\Phi}} \boldsymbol{\alpha}$ is consistent for $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}$. Consistent estimation of $\boldsymbol{\alpha}$ follows by transforming the estimate of $\boldsymbol{\alpha}^*$ by $\tilde{\mathbf{\Phi}}$.

Feasible Estimation of the Partially Restricted System

For our 171 Indonesian rice farms observed over six periods, if we can assume that $\mathbf{M}_1 = \ldots = \mathbf{M}_6 = \mathbf{M}$, $\rho_1 = \rho_3 = \rho_5 = \rho_W$, $\rho_2 = \rho_4 = \rho_6 = \rho_D$, $\sigma_1^2 = \sigma_3^2 = \sigma_5^2 = \sigma_W^2$, and $\sigma_2^2 = \sigma_4^2 = \sigma_6^2 = \sigma_D^2$, then we can impose the assumptions in equation (6) and estimate $\hat{\rho}_t$ and $\hat{\sigma}_t^2$, $t = 1, \ldots, 6$ as above. Then consistent estimates of ρ_W, ρ_D, σ_W^2, and σ_D^2 are the *partially restricted average estimates*

(11) $\hat{\rho}_W = \frac{1}{3}(\hat{\rho}_1 + \hat{\rho}_3 + \hat{\rho}_5)$, $\qquad \hat{\rho}_D = \frac{1}{3}(\hat{\rho}_2 + \hat{\rho}_4 + \hat{\rho}_6)$

and

$$\hat{\sigma}_W^2 = \tfrac{1}{3}(\hat{\sigma}_1^2 + \hat{\sigma}_3^2 + \hat{\sigma}_5^2), \qquad \hat{\sigma}_D^2 = \tfrac{1}{3}(\hat{\sigma}_2^2 + \hat{\sigma}_4^2 + \hat{\sigma}_6^2).$$

Again, these are two-stage estimates, which imply that

$$\hat{\mathbf{\Phi}}_W = (\mathbf{I}_N - \hat{\rho}_W \mathbf{M}) / \sqrt{\hat{\sigma}_W^2}\ , \quad \hat{\mathbf{\Phi}}_D = (\mathbf{I}_N - \hat{\rho}_D \mathbf{M}) / \sqrt{\hat{\sigma}_D^2}\ ,$$

which can be substituted into equation (5). Then fixed effect estimation of equation (5) with

$$\boldsymbol{\alpha}^* = [\ (\hat{\mathbf{\Phi}}_W \boldsymbol{\alpha})'\ \ (\hat{\mathbf{\Phi}}_D \boldsymbol{\alpha})'\ (\hat{\mathbf{\Phi}}_W \boldsymbol{\alpha})'\ \ (\hat{\mathbf{\Phi}}_D \boldsymbol{\alpha})'\ (\hat{\mathbf{\Phi}}_W \boldsymbol{\alpha})'\ \ (\hat{\mathbf{\Phi}}_D \boldsymbol{\alpha})'\ ]',$$

is consistent for $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}$.

Define $\varepsilon_W' = [\varepsilon_1'\ \varepsilon_3'\ \varepsilon_5']$, $\varepsilon_D' = [\varepsilon_2'\ \varepsilon_4'\ \varepsilon_6']$, $\tilde{\mathbf{u}}_W' = [\tilde{\mathbf{u}}_1'\ \tilde{\mathbf{u}}_3'\ \tilde{\mathbf{u}}_5']$ and $\tilde{\mathbf{u}}_D' = [\tilde{\mathbf{u}}_2'\ \tilde{\mathbf{u}}_4'\ \tilde{\mathbf{u}}_6']$.

Additionally, let $\bar{\tilde{\mathbf{u}}}_j = \mathbf{M} \tilde{\mathbf{u}}_j$, $\bar{\bar{\tilde{\mathbf{u}}}}_j = \mathbf{M}\ \bar{\tilde{\mathbf{u}}}_j$, $\bar{\boldsymbol{\varepsilon}}_j = \mathbf{M}\ \varepsilon_j$, and $\bar{\bar{\boldsymbol{\varepsilon}}}_j = \mathbf{M}\ \bar{\boldsymbol{\varepsilon}}_j$, $j = W, D$.

It follows analogously that the single-stage estimates are

(12) $$(\tilde{\rho}_j, \tilde{\sigma}_j^2) = \arg\min_{r,s^2} [\mathbf{v}_j(r,s^2)'\mathbf{v}_j(r,s^2) : s^2 \geq 0],\ j = W, D,$$

where

$$\mathbf{v}_j(\rho_j, \sigma_j^2) = \mathbf{G}_j[\rho_j \quad \rho_j^2 \quad \sigma_j^2]' - \mathbf{g}_j,\ j = W, D,$$

and where $\mathbf{G}_j$ and $\mathbf{g}_j$ are $\mathbf{G}_t$ and $\mathbf{g}_t$ of equation (6), but with j substituted for t and $3N$ substituted for N. Call these estimates the *partially restricted moment estimates*. The $\tilde{\rho}_j$ and $\tilde{\sigma}_j^2$ imply $\tilde{\mathbf{\Phi}}_j$ for wet and dry seasons, and fixed effect estimation of equation (5) is again consistent for $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}$.

So to summarize, the unrestricted estimation procedure yields $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ by solving equation (7); this is simply the application of the Kelejian and Prucha procedure T times. These estimates imply fully restricted average estimates ($\hat{\rho}$ and $\hat{\sigma}^2$) by averaging over T in equation (8) or partially restricted average estimates ($\hat{\rho}_j$ and $\hat{\sigma}_j^2$, j = W, D) by averaging over wet seasons and dry seasons in equation (11). These are two-stage estimates. Fully restricted moment estimates ($\tilde{\rho}$ and $\tilde{\sigma}^2$) are produced by solving equation (10), and partially restricted moments estimates ($\tilde{\rho}_j$ and $\tilde{\sigma}_j^2$, j = W, D) are produced by solving equation (12). These are single-stage estimates.

Application to Indonesian Rice Farms

We now estimate the models with a balanced panel of Indonesian rice farms. The data were previously analyzed by, e.g., Erwidodo; Lee and Schmidt; and Horrace and Schmidt (1996, 2000). For detailed discussion of the data see Erwidodo. For the panel specification of a stochastic frontier model, y is the natural logarithm of output (ln(rice)),

x is a vector of inputs (e.g., seed and fertilizer), and α_i embodies farm-level technical inefficiency. This is a standard stochastic frontier specification based on a Cobb-Douglas production function. Per Schmidt and Sickles, a measure of technical efficiency for farm i is calculated by plugging estimates of the α_i into the expression: $TE_i = \exp(\alpha_i - \max_j \alpha_j)$. In order to perform the spatial analysis we first specify the spatial weighting matrix, $\mathbf{M}_t$ for the error process, which captures productivity spillovers across farms.

Geographical and Climatic Characteristics of West Java

In 1977 the Indonesian Ministry of Agriculture began to survey 171 rice farms concerning farming practices over six (three wet and three dry) growing seasons. The farms were selected from six villages located in the production area of the Cimanuk River Basin in West Java. Of the six villages included in the sample, two are on the north coast of the island in an area with average altitudes of 10-15 meters above sea level. Another three villages are in an area (600-1100 meters above sea level) in the central part of West Java. The last village is in the center of the island with an average altitude of 375 meters. The infrastructure in the Cimanuk River Basin is fairly heterogeneous. Some of the villages (in both high and lowland areas) lack reliable transportation systems, and local roads are almost impassable in the wet (rainy) season. Other villages, located in close proximity to province capital cities, are highly accessible along paved, all-weather roads.[6]

Based on these facts, we construct and perform our analysis using two different weighting matrixes $\mathbf{M1}_t$ and $\mathbf{M2}_t$. The first one, $\mathbf{M1}_t$, is based on the *inverse of geographical distance* between individual farms.[7] We use geographical coordinates of the villages to determine physical distances between producing units. Distances between individual villages are between 31 and 91 km. The individual distances between farms within the same village are unavailable and are therefore arbitrarily chosen to be 10 km.[8] The $\mathbf{M1}_t$ weighting matrix then consists of the inverse values of these distances. That is, m_{ijt} equals the inverse of the distance between farms i and j. In the second weighting matrix we employ an intra-village contiguity scheme.[9] For $\mathbf{M2}_t$ we let m_{ijt} equal 1 if farms i and j are in the same village, equal 0 otherwise. That is, the weighting scheme is based on *common villages*. For computational simplification and as a standard practice in forming weighting matrixes, we *normalize* each weighting matrix so the elements of each row sum to one. Additionally, both the weighting schemes are assumed time invariant, so the t subscript can be dropped.

Spatial Analysis of Indonesian Rice Farms

We first estimate the standard fixed effect model of the stochastic production frontier described by equation (1). Inputs to the production of rice included in the data set are seed (kg), urea (kg), trisodium phosphate (TSP) (kg), labor (labor-hours), and land (hectares). Output is measured in kilograms of rice. The data also include dummy variables. *DP* equals 1 if pesticides were used and 0 otherwise. *DV*1 equals 1 if high yield varieties of rice were planted, and *DV*2 equals 1 if mixed varieties were planted. *DSS* equals 1 if it was a wet season, 0 otherwise. Results are in column I of Table 1 and are based on the restriction that $\rho_1 = ... = \rho_6 = 0$. These results are identical to those in Horrace and Schmidt (1996).

Before embarking on a spatial analysis, we use the residuals from the standard fixed effect estimation to determine whether or not spatial dependence (based on both weighting schemes) exists in the data. As before, let the usual fixed effect residuals in period t be $\tilde{\mathbf{u}}_t$. We employ two tests for spatial dependence; the first is the *Moran I* *statistic* (see, e.g., Anselin). (To preclude confusion with the symbol for the identity matrix we adopt the script ϑ.) The ϑ statistic for period t is

$$\vartheta_t = [N/S]\{[\tilde{\mathbf{u}}_t'\mathbf{M}\tilde{\mathbf{u}}_t]/\tilde{\mathbf{u}}_t'\tilde{\mathbf{u}}_t\},$$

where N is the number of farms and S is the sum of all elements in weighting matrix **M**. The null hypothesis for this test is "absence of spatial dependence."[10] Notice that we have dropped the t subscript on the weighting matrix **M**, because our empirical analysis assumes time invariance for this matrix. As shown by Cliff and Ord (1972) the asymptotic distribution for the statistic is standard normal, if ϑ is transformed in the usual manner

$$z_t = \{\vartheta_t - E[\vartheta_t]\}/V[\vartheta_t]^{1/2},$$

where $E[\vartheta_t]$ is the mean and $V[\vartheta_t]$ is the variance of the statistic in period t, derived under the null of "no spatial dependence". In the general case of a non-normalized weighting matrix these can be expressed in the form

$$E[\vartheta_t] = (N/S)tr(\mathbf{P}\,\mathbf{M})/(N-k),$$

$$V[\vartheta_t] = (N/S)^2\{tr(\mathbf{P}\,\mathbf{M}\,\mathbf{P}\mathbf{M}') + tr(\mathbf{P}\mathbf{M})^2 + [tr(\mathbf{P}\mathbf{M})]^2/(N-k)(N-k+2) - \{E[\vartheta_t]\}^2,$$

where **P** is the projection matrix $\mathbf{I}_N - \mathbf{x}_t(\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'$ and $\mathbf{x}_t$ is a matrix of the demeaned exogenous variables from the standard model in equation (1). The test is conducted for both weighting schemes in each time period t = 1, ..., 6. The z_t -scores for weighting scheme **M1** are in the third row (z_t) of Table 2 and range from 6.0702 in period t = 2 to 26.4159 in period t = 4. It is therefore safe to conclude that at the 95 percent confidence level we reject the hypothesis of "no spatial dependence" based on weighting scheme **M1**. Test results for weighting scheme **M2** were similar and are in the third row (z_t) of Table 4.

The Moran I statistic is sensitive to heteroskedasticity and tends to over-reject the null hypothesis when compared to the standard normal critical value. An alternative LM test procedure for the null hypothesis of no spatial dependence is presented by Anselin, Bera, Florax and Yoon (equation 9). The test statistic

$$LM_t = \frac{[\tilde{\mathbf{u}}_t' \mathbf{M} \tilde{\mathbf{u}}_t / \hat{\sigma}_t^2]^2}{tr[(\mathbf{M}' + \mathbf{M})\mathbf{M}]}$$

is distributed χ_1^2 with critical values of 3.84 (95 percent level) and 6.63 (99 percent level). Results are in the last row of Tables 2 and 4 for weighting schemes **M1** and **M2**, and confirm the Moran I results: we reject the null in each case.

Based on these test results, our proposed weighting schemes appear justified in each period. Consequently, we estimate the unrestricted spatial autoregressive parameters and error variances for each period for each scheme, using equation (6). Estimation results are in Tables 2 and 4 for **M1** and **M2**, respectively. Note that for both weighting schemes, the ρ-parameter tends to be larger in period 1 than in period 2, larger in period 3 than period 4, and larger in period 5 than in period 6. These differences correspond to differences in wet seasons (t = 1, 3, 5) and dry seasons (t = 2, 4, 6).

To identify parameter estimates for $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}$ in equation (5) we estimate the fully and partially restricted systems described in the last section. The fully restricted system, $\rho_1 = \ldots = \rho_6 = \rho$, is estimated using both the average autoregressive parameter of equation (8), $\hat{\rho}$, and the moments autoregressive parameter of equation (10), $\tilde{\rho}$, for each weighting scheme. Estimates of $\hat{\rho}$ = 0.7248 and $\tilde{\rho}$ = 1.0557 using weighting scheme **M1** are in Table 1, columns II and III. There is little difference in the slope parameter estimates based on $\hat{\rho}$ or $\tilde{\rho}$ or the standard FE model of column I. This is not surprising, since ignoring the spatial dependence causes an efficiency loss in the slope parameter estimates (not a bias). Indeed, the most noticeable differences in the estimates of columns I, II and III are in the standard error estimates, with columns II and III being generally smaller than column I, the standard model. The sign of the coefficient on the pesticide variable (DP) changes from positive to negative when we include spatial effects; however, it is always insignificant. The difference in the magnitudes of $\hat{\rho}$ and $\tilde{\rho}$ is troublesome. Perhaps this difference indicates that the restriction $\rho_1 = \ldots = \rho_6 = \rho$ does not hold. We did not attempt to test this, however it would be possible if the variance matrix of the ρ_t were estimable. The results of the fully restricted model under weighting scheme **M2** are in columns II and III in Table 3. The results are similar to the **M1** case: slope coefficients do not change much, standard error estimates decrease, and there is a large difference between the two estimates of ρ.

Feasible estimation of the partially restricted system follows the same pattern, except that instead of only one correlation coefficient fixed for all time periods now we estimate and utilize two correlation coefficients--one for the wet and one for the dry season. We calculate the average parameter estimates of equation (11), $\hat{\rho}_W$ and $\hat{\rho}_D$, and the moments estimates of equation (12), $\tilde{\rho}_W$ and $\tilde{\rho}_D$, for each weighting scheme. Fixed effect estimation results for ($\hat{\rho}_W$, $\hat{\rho}_D$) and ($\tilde{\rho}_W$, $\tilde{\rho}_D$), based on weighting scheme **M1**, are in Table 1, columns IV and V. The differences between the average and moments parameter estimates are much less pronounced than in the fully restricted case (compare estimate $\hat{\rho}_W$ = 0.7584 to $\tilde{\rho}_W$ = 0.8218, estimate $\hat{\rho}_D$ = 0.6914 to $\tilde{\rho}_D$ = 0.7476, and estimate $\hat{\rho}$ = 0.7248 to $\tilde{\rho}$ = 1.0557). One might conclude that the partially restricted model seems to fit the data better; however, this is not formally tested. (Additionally, the fact that the estimates are now all less than unity suggests that the partially restricted model may be favored over the fully restricted model.) Again, the standard errors of the slope parameter estimates are smaller for the partially restricted model than for the standard model (column I). The coefficient on the season variable (DSS) is not identified, since it is effectively time invariant now that the data set has been dichotomized into "wet" and "dry" subsamples.[11] The coefficients on the partially restricted system are generally higher than those of the fully restricted system (columns II and III) and the standard model (column I). As in the fully restricted system, the coefficient on the pesticide variable (DP) is negative and insignificant. Even though it is insignificant, this is troubling, since economic theory usually dictates that a production function be non-decreasing in its arguments. However, one could argue that too much pesticide might have a negative effect on output. Alternatively, one could argue that we have not adequately controlled for the interaction between pesticides (DP), output (y) and weather (DSS, ρ_W and ρ_D). Perhaps pesticide use is higher during the wet season (more water, more insects), and our simple dummy variable for pesticide does not adequately capture a more complex relationship. Nonetheless, the implications are compelling and the coefficient *is* insignificant. Estimation results for weighting scheme **M2** are similarly presented in columns IV and V of Table 3. Again, results are similar to scheme **M1** for this particular sample.

Technical Efficiency Rankings

Stochastic frontier analyses are often concerned with estimating firm-level technical inefficiency and, in particular, determining the relative magnitudes of the resulting inefficiency measures, using a rank or order statistic. In the next analysis we demonstrate how the various weighting schemes affect the technical efficiency rankings of the farms. Specifically, for each weighting scheme we estimate and rank the estimated technical efficiencies, $\exp(\alpha_i - \max_j \alpha_j)$, for each farm. This is done for the standard fixed effect model (column I of Table 1) and for *the fully restricted moments estimator* (column III of Tables 1, and 3). The idea is to see how the rankings differ between the standard model and the spatial model for both of the weighting schemes. Order statistics for each model are in Table 5. The first three columns of the table are results for the standard fixed effect model. Since there are 171 farms we only report results for the 4 farms with the highest technical efficiency, the 4 farms with the median technical efficiency, and the 4 farms with the lowest technical efficiency. Column I contains the farm number, column II contains the ordered estimates of farm-level technical efficiency, and column III contains the ordinal rankings for the standard fixed effect model (numbered 1 to 171). To see the effects of spatial dependence on the technical efficiency estimation, we also report the ordinal rankings for the same 12 farms for the fully restricted spatial model under weighting schemes **M1** and **M2** in columns IV and V. While there are some changes across weighting schemes in the rank ordering of the most and least efficient farms, these are minor. For instance, in the standard model, farm 152 had a technical efficiency rank of 4, but it has a rank of 6 under the weighting schemes. Notice that the ranking of the most efficient farm (farm 164) is always 1 and that of the least efficient farm (farm 45) is always 171. The largest differences in ranking appear in the median farms. For example, farm 166 has a standard fixed effect ranking of 85 but spatial rankings of 116. These are potentially large changes in the median technical efficiency ranking, which could only be detected with a spatial analysis.

To further summarize the changes in the efficiency ranking in Table 5, we calculate *Spearman's rho* (r_s) for both weighting schemes, using the standard fixed effect model as the baseline. Spearman's rho is a standard measure of rank correlation between two rank statistics given by

$$r_s = 1 - \frac{6\sum_i \delta_i^2}{N^3 - N},$$

where δ_i is the difference in the rankings for the i^{th} farm. For example when comparing the rank statistic for the standard model and the **M1** model in Table 5, $\delta_{164} = 0$ and $\delta_{15} = 86 - 62 = 24$. Here we always compare the rankings of the **M1** and **M2** models to the standard model ranking. It is true that $r_s \in [-1, 1]$, $r_s = 1$ when the two rank statistics are identical, and $r_s = -1$ when the rank statistics are completely reversed (i.e., as we move from one order statistic to the other, the most efficient farm becomes the least efficient, the second most efficient farm become the second-least efficient...). Spearman statistics are in the last row of Table 5 and are on the order of 0.8 for both of the weighting schemes. We can interpret this result as saying that only 80 percent of the rank statistic is preserved when we use a spatial weighting specification over the standard specification.

To better understand the changes in technical efficiency under the various weighting schemes, we present some density plots of the estimates of the parameters, α_i. Technically, there is no distribution of α_i to speak of, since it is assumed to be a fixed parameter and not a random variable. The *estimates* of the α_i are indeed random, and each estimate has its own marginal distribution from the joint distribution of the estimate of the N-dimensional vector, α. However, for the purposes of exposition, we treat the estimates of the α_i as if they are random draws from a univariate distribution in what follows. According to the panel data specification of Schmidt and Sickles, $\alpha_i = \alpha^{max} - \tau_i$, where τ_i is the non-negative technical efficiency of farm i and α^{max} is a parameter representing maximal efficiency. The implication is that for fixed α^{max}, the "distribution" of α_i is just a relocation of the "distribution" of technical inefficiency. Therefore to infer the effects of various weighting schemes on the estimates of technical efficiency is to make inferences on the estimates of α_i.

Density plots for the various models are in Figure 1. Density estimates are based on maximum likelihood, cross-validation bandwidth selection and a standard Gaussian kernel. Fixing α^{max} across models, some generalizations about this data set can be made. First, the standard fixed effect model (FE in the figure) without spatial lags in the errors tends to underestimate α_i (overestimate technical inefficiency) in comparison to the spatial models (**M1** and **M2**). This is technical inefficiency in an absolute sense, since we are fixing α^{max} across models at some unknown value. This is reflected in Figure 1 as the density of the standard fixed effect model (FE) being shifted to the left of the densities for the spatial models (with little to no rescaling). This has implications for predictions of the conditional mean output implied by equation (1): the fixed effect model (on average) gives lower predictions of productive output than the spatial models (all things being equal). That is, for fixed technology and input factors, the spatial models impute more of the observed output to unobserved technical ability (α_i) and less of it to luck (u_{it}) in this data set. Indonesian rice farms may be operating closer to the efficient frontier than previous studies suggest.

Conclusions

This paper has presented a generalization of the cross-sectional model of Kelejian and Prucha. Because economic agents and entities have finite lives, one cannot always rely on large T in economic panel data sets. Most panel data sets (with the exception of perhaps microeconomic financial data) have large N and small T. Additionally, if T is somewhat large, the usually time-invariant unobserved heterogeneity models (e.g., fixed effect) may not be applicable, since it is widely held that heterogeneity may change in long-run, dynamic economic systems (particularly when it is viewed as technical inefficiency). The result is that consistency arguments usually must hinge on N asymptotics. This is fine for estimating conditional means (the model's slope parameters). However, any second moment parameters that embody cross-sectional dependence cannot be consistently estimated in the sense that they will necessarily rely on T asymptotics.

When faced with this dilemma, researchers have two recourses: collect more data or impose more structure on the model and hope that the structure will be testable. Given the aforementioned arguments against large T, it would seem that we are faced with the alternative of imposing more structure on our models. The question then becomes: what structure is reasonable? Spatial weighting schemes seem to be a reasonable and natural approach. The theoretical economic literature is rife with arguments for economic spillovers, and spatial analysis provides a means to make these spillovers explicit. Moreover, tests of "no spatial dependence" do exist in the literature. Therefore, if we must make assumptions about the second moments of our data, spatial weighting schemes may be a viable approach.

Dynamic spatial dependence in the second moment of our estimators has implications for dynamics in the first moment. The fixed effect model has time-invariant heterogeneity parameters, but the transformed model had dynamic parameters. It is this loss of time-invariance that makes the general model "not identified", and forces us to impose some restrictions on the dynamics of the spatial dependence. This could be important. Most panel data models that attempt to make the heterogeneity parameters dynamic do so by imposing structure on the first moments of the models. For instance, several papers in the stochastic frontier literature impose special structure on the conditional mean of the heterogeneity parameters. (For examples see Cornwell, Schmidt and Sickles; Lee and Schmidt; Battese and Coelli; and Kumbhakar.) The models presented here create dynamic heterogeneity through *second moment* conditions on the error process. The implications of this difference for models of dynamic heterogeneity are unknown, but it is interesting to point this difference out.

Additionally, spatial dependence may be a way to indirectly incorporate time-invariant regressors into a fixed effect model. For example, Horrace and Schmidt (1996) incorporate dummy variables for the six villages into a random effects specification but are forced to exclude these dummy variables from a fixed effect specification, because they are time-invariant at the farm level. In the application presented here, village effects are incorporated into the second moment of the residual for the fixed effects model. While there are commonly employed techniques for incorporating time-invariant regressors into a fixed effect model (see Hausman and Taylor), the research presented here provides analysts with an alternative means of accomplishing this.

Table 1. Rice Regressions, Weighting Scheme **M1** – Inverse of Distance

	Standard FE Model I	Fully Restricted Average II	Fully Restricted Moment III	Partially Restricted Average IV	Partially Restricted Moment V
$\hat{\rho}$	-	0.7248	-	-	-
$\tilde{\rho}$	-	-	1.0557	-	-
$\hat{\rho}_w$	-	-	-	0.7584	-
$\hat{\rho}_D$	-	-	-	0.6914	-
$\tilde{\rho}_W$	-	-	-	-	0.8218
$\tilde{\rho}_D$	-	-	-	-	0.7476
Seed	0.1208* (0.030)	0.1038* (0.025)	0.0998* (0.024)	0.1292* (0.024)	0.1248* (0.024)
Urea	0.0918* (0.021)	0.0894* (0.018)	0.0901* (0.017)	0.1405* (0.015)	0.1440* (0.015)
TSP	0.0892* (0.013)	0.0353* (0.012)	0.0244* (0.012)	0.0340* (0.011)	0.0307* (0.011)
Labor	0.2431* (0.032)	0.2366* (0.029)	0.2379* (0.028)	0.2254* (0.026)	0.2204* (0.026)
Land	0.4521* (0.035)	0.4879* (0.031)	0.4931* (0.030)	0.5046* (0.027)	0.5141* (0.027)
DP	0.0338 (0.032)	-0.0178 (0.028)	-0.0298 (0.028)	-0.0224 (0.025)	-0.0212 (0.025)
DV1	0.1788* (0.041)	0.1084* (0.038)	0.0935* (0.038)	0.1250* (0.034)	0.1320* (0.035)
DV2	0.1754* (0.057)	0.1060* (0.049)	0.0952* (0.048)	0.0917* (0.048)	0.0947* (0.048)
DSS	0.0533* (0.022)	0.0759 (0.063)	0.1062 (0.302)	- -	- -
R^2	0.9102	0.9246	0.9271	0.9190	0.9177

Notes: Numbers in parentheses are standard errors.
* slope estimate significant at 5% level.

Table 2. Unrestricted Estimates of ρ and σ^2, and Tests of Weighting Scheme **M1**

Time period	1	2	3	4	5	6
$\hat{\rho}_t$	0.62	0.52	0.87	0.84	0.77	0.71
$\hat{\sigma}_t^{\,2}$	0.04	0.08	0.08	0.07	0.05	0.07
z_t	8.24	6.07	24.95	26.42	14.19	12.30
LM_t	65.41	30.50	1461.01	1680.70	254.85	175.99

Table 3. Rice Regressions, Weighting Scheme **M2** – Common Villages

	Standard FE Model I	Fully Restricted Average II	Fully Restricted Moment III	Partially Restricted Average IV	Partially Restricted Moment V
$\hat{\rho}$	-	0.6604	-	-	-
$\tilde{\rho}$	-	-	0.9882	-	-
$\hat{\rho}_w$	-	-	-	0.6811	-
$\hat{\rho}_D$	-	-	-	0.6398	-
$\tilde{\rho}_W$	-	-	-	-	0.7388
$\tilde{\rho}_D$	-	-	-	-	0.6999
Seed	0.1208*	0.1035*	0.0996*	0.1255*	0.1248*
	(0.030)	(0.025)	(0.024)	(0.024)	(0.024)
Urea	0.0918*	0.0909*	0.0901*	0.1435*	0.1446*
	(0.021)	(0.018)	(0.017)	(0.015)	(0.015)
TSP	0.0892*	0.0356*	0.0239*	0.0326*	0.0301*
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)
Labor	0.2431*	0.2385*	0.2376*	0.2201*	0.2198*
	(0.032)	(0.029)	(0.028)	(0.026)	(0.026)
Land	0.4521*	0.4855*	0.4934*	0.5131*	0.5148*
	(0.035)	(0.031)	(0.030)	(0.028)	(0.027)
DP	0.0338	-0.0189	-0.0306	-0.0208	-0.0219
	(0.032)	(0.028)	(0.028)	(0.025)	(0.025)
DV1	0.1788*	0.1116*	0.0928*	0.1335*	0.1326*
	(0.041)	(0.038)	(0.038)	(0.034)	(0.035)
DV2	0.1754*	0.1080*	0.0947*	0.0970*	0.0961*
	(0.057)	(0.049)	(0.048)	(0.049)	(0.049)
DSS	0.0533*	0.0789	0.0844	-	-
	(0.022)	(0.051)	(1.424)	-	-
R^2	0.9102	0.9240	0.9271	0.9171	0.9174

Notes: Numbers in parentheses are standard errors.
* slope estimate significant at 5% level.

Table 4. Unrestricted Estimates of ρ and σ^2, and Tests of Weighting Scheme **M2**

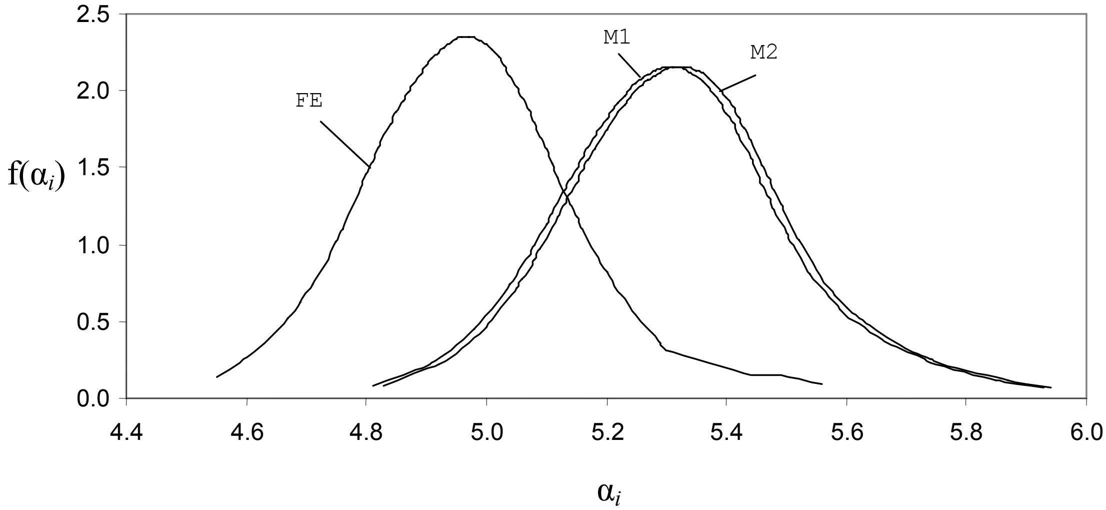
Time period	1	2	3	4	5	6
$\hat{\rho}_t$	0.57	0.48	0.79	0.79	0.69	0.65
$\hat{\sigma}_t^{\ 2}$	0.04	0.08	0.08	0.07	0.05	0.07
z_t	7.50	5.09	23.54	23.51	13.67	11.18
LM_t	57.03	21.62	1409.13	1386.30	256.73	153.03

Table 5. Technical Efficiency Orders Statistics, Various Models

Standard FE Model			Spatial Models	
Farm # I	Standard FE Efficiency II	Standard FE Model III	Weight Scheme **M1** IV	Weight Scheme **M2** V
164	100%	1	1	1
118	93.23%	2	2	2
163	93.03%	3	3	3
152	89.93%	4	6	6
13	55.62%	84	106	106
166	55.47%	85	116	116
15	55.40%	86	62	62
40	55.35%	87	54	54
86	39.80%	168	165	165
143	38.37%	169	169	169
117	37.90%	170	168	168
45	36.55%	171	171	171
	r_s:	1.0000	0.8027	0.8095

Notes: Spatial results are for the fully restricted moments estimator.
Technical efficiency for farm $i = \exp\{\alpha_i - \max_j \alpha_j\}$.

Figure 1. Density Estimates of α_i for Various Models



FE = fixed effect with no spatial weighting
M1 = M1 weighting scheme
M2 = M2 weighting scheme

Appendix

Assumptions 3, 4, and 5 from Kelejian and Prucha (1999). Let $\mathbf{P}(\rho_t) = (\mathbf{I}_N - \rho_t\mathbf{M}_t)^{-1}$ with typical element $p_{ij}(\rho_t)$.

Assumption 3: (*i*) The sums $\Sigma_i|m_{ijt}|$ and $\Sigma_j|m_{ijt}|$ are bounded by, say, $c_m < \infty$ for all $1 \leq i, j \leq N$, $N \geq 1$. (*ii*) The sums $\Sigma_i| p_{ij}(\rho_t)|$ and $\Sigma_j| p_{ij}(\rho_t)|$ are bounded by, say, $c_p < \infty$ for all $1 \leq i, j \leq N$, $N \geq 1$, $|\rho_t| < 1$.

Assumption 4: Let $\tilde{u}_{it}$ be the i^{th} element of $\tilde{\mathbf{u}}_t$. There exists finite dimensional random vectors $\mathbf{d}_{it}$ and Δ_t such that $|\tilde{u}_{it} - u_{it}| \leq ||\mathbf{d}_{it}||\ ||\Delta_t||$ with $N^{-1}\Sigma_i||\mathbf{d}_{it}||^{2+\delta} = O_p(1)$ for some $\delta > 0$ and $N^{-1/2}\Sigma_i||\Delta_t|| = O_p(1)$.

Assumption 5: The smallest eigenvalue of $\Gamma_t'\Gamma_t$ is bounded away from zero.

References

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

Anselin, L. *Spatial Econometrics: Methods and Models*. London and Dordrecht: Kluwer Academic, 1988.

Anselin, L., A.K. Bera, R. Florax, and M.J. Yoon. "Simple Diagnostic Tests for Spatial Dependence." *Regional Science and Urban Economics* 26 (February 1996): 77-104.

Battese, G.E., and T.J. Coelli. "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India." *Journal of Productivity Analysis* 3 (1992): 149-165.

Case, A.C. "Spatial Patterns in Household Demand." *Econometrica* 59 (July 1991): 953-965.

Cliff, A.D., and J.K. Ord. *Spatial Autocorrelation*. Volume 5, *Monographs in Spatial and Environmental Systems Analysis*. London: Pion, 1973.

______. "Testing for Spatial Autocorrelation among Regression Residuals." *Geographical Analysis* 4 (1972): 268-284.

Conley, T.G. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (September 1999): 1-45.

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

De Long, J.B., and L.H. Summers. "Equipment Investment and Economic Growth." *Quarterly Journal of Economics* 106 (May 1991): 445-502.

Dubin, R.A. "Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms." *Review of Economics and Statistics* 70 (August 1988): 466-474.

Erwidodo. "Panel Data Analysis on Farm-Level Efficiency, Input Demand and Output Supply of Rice Farming in West Java, Indonesia." Ph.D. Dissertation. Department of Agricultural Economics, Michigan State University, 1990.

Fishback, Price V., William C. Horrace, and Shawn Kantor. "Federal Programs in Times of Crisis: The Impact of the New Deal on Local Economies During the Great Depression." Department of Economics Working Paper. Tucson: University of Arizona, 2002.

Geary, R. "The Contiguity Ratio and Statistical Mapping." *The Incorporated Statistician* 5(1954): 115-145.

Hausman, J.A., and W.E. Taylor. "Panel Data and Unobservable Individual Effects." *Econometrica* 49 (November 1981): 1377-1398.

Horrace, W.C., and P. Schmidt. "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models." *Journal of Productivity Analysis* 7 (July 1996): 257-282.

______. "Multiple Comparisons with the Best, with Economic Applications." *Journal of Applied Econometrics* 15 (January-February 2000): 1-26.

Kelejian, H.H., and I.R. Prucha. "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model." *International Economic Review* 40 (May 1999): 509-533.

Kelejian, H.H., and D.P. Robinson. "A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a County Expenditure Model." *Papers in Regional Science* 72(1993): 297-312.

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

Lee, Y.H., and P. Schmidt. "A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency." *The Measurement of Productive Efficiency: Techniques and Applications.* H.O. Fried, C.A.K. Lovell, and S.S. Schmidt, eds., pp. 237-255. New York, Oxford: Oxford University Press, 1993.

Meeusen, W., and J. van den Broeck. "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *International Economic Review* 18 (June 1977): 435-444.

Moran, P. "The Interpretation of Statistical Maps." *Journal of the Royal Statistical Society (B)* 10(1948): 243-251.

Moulton, B.R. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit." *Review of Economics and Statistics* 72 (May 1990): 334-338.

Quah, D. "International Patterns of Growth: I. Persistence in Crosscountry Disparities." Working Paper. London: London School of Economics. October, 1992.

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

Topa, G. "Social Interactions, Local Spillovers and Unemployment." Ph.D. Dissertation. University of Chicago, 1996.

Notes

---

1 The Conley technique could also be applied here and could conceivably produce more flexible results since Conley's technique accommodates less restrictive assumptions on the error process. However, our intent here is to specifically examine the Kelejian and Prucha results.

2 We present no proofs of our results, because they are all straightforward extensions of Kelejian and Prucha's proofs.

3 Assumptions 3, 4, and 5 of Kelejian and Prucha are contained in the Appendix.

4 The fact that we estimate ρ_t, $t = 1, ..., T$, implies a test of the hypothesis $\rho_1 = \ldots = \rho_T = \rho$. We are not aware of any such test, nor are we aware of a standard error calculation for the estimate of ρ_t. Of course, the standard error could be bootstrapped. Later, we use the Moran *I* test and the Lagrange Multiplier test to test the significance of the overall weighting scheme in each period.

5 Notice that the middle moment condition contains N^{-1} and not $(TN)^{-1}$, since it is based on **M** and not $\mathbf{M}^*$.

6 The survey ended in 1983, so the infrastructure description may be different from the current state.

7 Cliff and Ord (1973) first measured potential interactions between spatial units using a combination of distance measures and relative length of the common border (contiguity). Since there is no true measure of contiguity available in our case we use physical distance only as a proxy for interdependence between spatial units.

8 Experimentation with the weighting matrix suggested that the analysis was fairly robust to this arbitrary selection.

[9] Both Moran and Geary advanced initial measures of spatial dependence (spatial correlation) that were based on the notion of "binary contiguity" between spatial units. That is, if spatial units have a common border (are contiguous) a value of 1 is assigned to the spatial correlation; 0 otherwise.

[10] In the words of Anselin, interpretation of the test is not always straightforward, even though it is by far the most widely used approach. Indeed, while the null hypothesis is obviously the absence of spatial dependence, a precise expression for the alternative does not exist.

[11] Even though the time dimension has effectively been cut in half by this dichotomy, the estimates of the slope parameters are still based on the entire sample (*TN*) after the observables have been demeaned, based on whether they are "dry" or "wet."