TECHNICAL WORKING PAPER SERIES

GENERALIZED MOMENTS ESTIMATION FOR SPATIAL PANEL DATA

Viliam Druska
William C. Horrace

Technical Working Paper 291
http://www.nber.org/papers/T0291

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2003

Generalized Moments Estimation for Panel Data
Viliam Druska and William C. Horrace
NBER Technical Working Paper No. 291
February 2003
JEL No. C21, C23, D24

ABSTRACT

This paper considers estimation of a panel data model with disturbances that are autocorrelated across cross-sectional units. It is assumed that the disturbances are spatially correlated, based on some geographic or economic proximity measure. If the time dimension of the data is large, feasible and efficient estimation proceeds by using the time dimension to estimate spatial dependence parameters. For the case where the time dimension is small (the usual panel data case), we develop a generalized moments estimation approach that is a straight-forward generalization of a cross-sectional model due to Kelejian and Prucha. We apply this approach in a stochastic frontier framework to a panel of Indonesian rice farms where spatial correlations are based on geographic proximity, altitude and weather. The correlations represent productivity shock spillovers across the rice farms in different villages on the island of Java. Test statistics indicate that productivity shock spillovers may exist in this (and perhaps other) data sets, and that these spillovers have effects on technical efficiency estimation and ranking.

Viliam Druska
CERGE
Charles University

William C. Horrace
Center for Policy Research
Syracuse University
426 Eggers Hall
Syracuse, NY 13244-1020
and NBER
whorrace@maxwell.syr.edu

Generalized Moments Estimation for Spatial Panel Data

Abstract:
This paper considers estimation of a panel data model with disturbances that are autocorrelated across cross-sectional units. It is assumed that the disturbances are *spatially correlated*, based on some geographic or economic proximity measure. If the time dimension of the data is large, feasible and efficient estimation proceeds by using the time dimension to estimate spatial dependence parameters. For the case where the time dimension is small (the usual panel data case), we develop a generalized moments estimation approach that is a straight-forward generalization of a cross-sectional model due to Kelejian and Prucha. We apply this approach in a stochastic frontier framework to a panel of Indonesian rice farms where spatial correlations are based on geographic proximity, altitude and weather. The correlations represent productivity shock spillovers across the rice farms in different villages on the island of Java. Test statistics indicate that productivity shock spillovers may exist in this (and perhaps other) data sets, and that these spillovers have effects on technical efficiency estimation and ranking.

1. Introduction

Over the last few years much has been written on the subject of the spatial dependence in cross-sectional economic data. These are data in which observations can be characterized by absolute or relative location, based on some form of coordinate system or distance measure. 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. This spatial dependence stems from the existence of implicit functional relations between units. Anselin provides an excellent textbook treatment of this phenomenon. 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 or economic spill-overs, to name a few.

In a recent article, Kelejian and Prucha consider generalized moments estimation of a certain class of these models which allows for spatial autocorrelation of econometric *disturbances* across cross-sectional units. This “spatial autocorrelation or dependence” can be likened to the autocorrelation theory of the time series literature, however in this case estimation hinges on the *ex ante* specification of a “spatial weighting matrix” in the cross-sectional dimension of the error process. The form of the weighting matrix is at the discretion of the analyst, but very often can be based on some underlying economic, geographic, of meteorological theory. The specification of the spatial weighting matrix may seem like a fairly strong parametric assumption to impose on the model, however the hypothesis of ‘no spatial dependence’ is testable and is the price paid for lack of a time dimension in the data. “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, p. 1). Unfortunately, in the usual panel data case, the cross-sectional dimension is large and the time dimension is *small* (fixed), so consistent estimation of the cross-sectional correlations is typically not justified.

This paper is concerned with generalized moments estimation of this class of models under normal panel data conditions. Specifically, we apply the Kelejian and Prucha estimator to the usual panel data case and introduce a generalization of their estimator based on certain restrictions on the evolution of the 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 more standard approaches that hinge on T asymptotics.

This paper also presents an empirical application of these spatial panel techniques to the efficiency measurement problem of the well-developed stochastic frontier literature, which attempts to model a common production function for a sample of firms based on observables (inputs and outputs) and an unobservable, one-sided component viewed as technical or production inefficiency. Cross-sectional estimation of these models is usually traced back 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 6 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 may not be justified. Consequently, we specify a spatial weighting scheme in the error process which allows for spillovers across farms based on geographic proximity, altitude and weather conditions. The results indicate that spatial correlations exist in the data and have ramifications for the estimation of the production function and the estimation of farm-level technical efficiency.

The paper is organized as follows. The next section presents an unrestricted, general panel data model based on the Kelejian and Prucha technique and also fully- and partially-restricted models that are used in the application. Section 3 discusses feasible estimation of these models. Section 4 presents the Indonesian rice farm example. Section 5 summarizes and concludes.

2.1 A Panel Model with Spatial Disturbances

Consider the standard fixed effect (FE) model:

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

where β is $(k\times 1)$ and x_{it} is $(1\times k)$. Here we assume that T is fixed, so we cannot rely on T-asymptotics. Collecting i the model becomes

(1) $y_t = \alpha + x_t\beta + u_t, \quad t = 1, ..., T,$

where $\alpha' = [\alpha_1, ..., \alpha_N]$ and x_t is $(N\times k)$. Now suppose that the error term is spatially lagged such that

(2) $u_t = \rho_t M_t u_t + \varepsilon_t; \quad t = 1, ..., T,$

where ρ_t is a scalar, spatial autoregressive parameter and M_t is a $(N\times N)$ spatial weighting matrix of known constants, which captures the spatial correlations across cross-sectional units. (In the sequel we allow for a time-invariant spatial parameter and weighting matrix.) Elements of M_t are m_{ijt} and are chosen based on some geographic or economic proximity measure such are contiguity or physical, economic or climatic distances or dissimilarities. For example, in section 4, 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. Traditional treatments of the stochastic frontier model hypothesize output as a linear function of a) farm-level technical (in)efficiency (an unobserved factor imputed to each farm), b) a representative log-production function (deterministic and within the control of each farm) and c) productivity shocks (random and out of the farmer's control). These three additive elements of output are captured in equation (1) by the matrices α, $x_t\beta$ and u_t , respectively. 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 experienced by other farms in the sample. This would seem reasonable if productivity shocks included geographic or climatic unobservables that effected 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 y_t on the right-hand side of equation (1). (Indeed, the spatial econometric literature addresses estimation of models with spatially lagged endogenous explanatory variables.) Therefore, the specification also *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 would seem reasonable if the production function is viewed as a purely deterministic (engineering) process, where the farmer controls all inputs. We will also have need of the additional assumptions:

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

> *Assumption 2:* All diagonal elements of M_t are zero. The matrix $(I_N - \rho_t M_t)$ is non-singular. $|\rho_t| < 1$.

Notice that under Assumptions 1 and 2, $u_t = (I_N - \rho_t M_t)^{-1}\varepsilon_t$, so $E(u_t) = 0$ for all t, but $E(u_t u_t')$ has a general, non-spherical structure, which is a function of ρ_t, M_t and σ_t^2. Since M_t is known, $E(u_t u_t')$ 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 $\rho_t = \rho$, $M_t = M$ and $\sigma_t^2 = \sigma^2$, then $E(u_t u_t')$ is a constant, which can be consistently estimated as $T \to \infty$. Here, we assume that T is fixed, so consistent estimation of E($u_t u_t'$) is unreasonable, and we must assume that M_t is known to identify an estimate of equation (1). For now, assume that ρ_t and σ_t^2 are known. Collecting t,

(3) $y = \iota_T \otimes \alpha + x\beta + u, \quad u = (\rho^* \otimes I_N) M^* u + \varepsilon,$

where ι_T is a T dimensional column vector of ones, x is ($TN \times k$) and

$$M^* = \begin{bmatrix} M_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & M_T \end{bmatrix} \quad , \quad \rho^* = \begin{bmatrix} \rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \rho_T \end{bmatrix}.$$

Notice that

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

so the disturbance in equation (2) is heteroskedastic. Define $\Phi_t = (I_N - \rho_t M_t)/\sigma_t$ then we can pre-multiply the model in equations (1) and (2) to get,

(4) $y_t^* = \alpha_t^* + x_t^* \beta + \varepsilon_t^*,$

where $y_t^* = \Phi_t y_t$, $x_t^* = \Phi_t x_t$, $\alpha_t^* = \Phi_t \alpha$ and $\varepsilon_t^* = \Phi_t \varepsilon_t = \varepsilon_t / \sigma_t$. Collecting t

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

where $\alpha^*{}' = [\alpha_1^*{}', ..., \alpha_T^*{}']$, a TN dimensional column vector. Equation (5) possess a "well-behaved" disturbance. That is, E(ε^*) = 0 and E($\varepsilon^* \varepsilon^*{}'$) = I_{TN}. Identification of any estimates of the parameters in equation (5) hinges on estimation of the unknown parameters M_t, ρ_t, σ_t^2, which will be ultimately undertaken in the sequel. The Kelejian and Prucha cross-sectional procedure could be directly applied to equation (4) T times over N observations to recover estimates of M_t, ρ_t, σ_t^2. These estimates could then be used to estimate the parameters in equation (5).[i] We refer to this estimation technique as "unrestricted estimation". However, our particular application implies some restrictions on the model in equation (5) that we can exploit to improve the efficiency of the estimation (as long as those restrictions hold in the sample). In particular, our definitions of spatial dependence are based on distinct physical characteristics of the farming villages on the island of Java (altitude, longitude, latitude, infrastructure, etc.) which are certainly constant over the short time period of the data (6 years). Therefore, it may be particularly useful to impose some additional structure on equation (5), and it is this additional structure that produces a modest but meaningful generalization of the Kelejian and Prucha results.

2.1 A Fully Restricted Specification.

One obvious restriction would be to assume that some subset of the weighting matrices, autoregressive parameters and variance parameters are equal. As an extreme case we could assume that $M_1 = \ldots = M_T = M$, $\rho_1 = \ldots = \rho_T = \rho$, $\sigma_1^2 = \ldots = \sigma_T^2 = \sigma^2$, implying $\Phi_1 = \ldots = \Phi_T = \Phi$. Then $\alpha_t^* = \Phi\alpha^*$ in equation (4) and $\alpha^* = \iota_T \otimes \Phi\alpha$ in equation (5). Of course, the error term ε of equation (3) is no longer heteroskedastic ; it has variance matrix $E(\varepsilon\varepsilon') = \sigma^2 I_{TN}$, so Φ 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 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 Φ. 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 section 3.

2.2 A Partially Restricted Specification.

As another example of a reasonable restriction on the parameters of the model, briefly consider the empirical example of the sequel. 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), similarly for M_t, σ_t^2 and Φ_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 spill-overs based on infrastructure are potentially diminished relative to the dry season.) 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 2×171 = 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.

3. 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.[ii] Using their notation, let $\tilde{u}_t$ be a predictor of u_t from the fixed effect (or within) regression implied by equation (1), ignoring equation (2). That is, $\tilde{u}_t$ converges in distribution to the random variable u_t . Additionally, let $\tilde{\bar{u}}_t = M_t \tilde{u}_t$, $\tilde{\bar{\bar{u}}}_t = M_t \tilde{\bar{u}}_t$, $\bar{\varepsilon}_t = M_t \varepsilon_t$, and $\bar{\bar{\varepsilon}}_t = M_t \bar{\varepsilon}_t$. Consider the following $3T$ moment conditions implied by equations (1) and (2) and assumptions 1 and 2.

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

$t = 1, ..., T$. Noting that $\varepsilon_t = (I_N - \rho_t M_t)u_t$, these moment conditions imply the following system of $3T$ equations,

$$\Gamma_t[\rho_t, \rho_t^2, \sigma_t^2]' - \gamma_t = 0,$$

where,

$$\Gamma_t = \begin{bmatrix} \frac{2}{N}E(u_t'\bar{u}_t) & \frac{-1}{N}E(\bar{u}_t'\bar{u}_t) & 1 \\ \frac{2}{N}E(\bar{\bar{u}}_t'\bar{u}_t) & \frac{-1}{N}E(\bar{\bar{u}}_t'\bar{\bar{u}}_t) & \frac{1}{N}tr(M_t'M_t) \\ \frac{1}{N}E(u_t'\bar{\bar{u}}_t + \bar{u}_t'\bar{u}_t) & \frac{-1}{N}E(\bar{u}_t'\bar{\bar{u}}_t) & 0 \end{bmatrix}, \qquad \gamma_t = \begin{bmatrix} \frac{1}{N}E(u_t'u_t) \\ \frac{1}{N}E(\bar{u}_t'\bar{u}_t) \\ \frac{1}{N}E(u_t'\bar{u}_t) \end{bmatrix},$$

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

(6) $$G_t[\rho_t, \rho_t^2, \sigma_t^2]' - g_t = \nu_t(\rho_t, \sigma_t^2),$$

$$G_t = \begin{bmatrix} \frac{2}{N}\tilde{u}_t'\tilde{\bar{u}}_t & \frac{-1}{N}\tilde{\bar{u}}_t'\tilde{\bar{u}}_t & 1 \\ \frac{2}{N}\tilde{\bar{\bar{u}}}_t'\tilde{\bar{u}}_t & \frac{-1}{N}\tilde{\bar{\bar{u}}}_t'\tilde{\bar{\bar{u}}}_t & \frac{1}{N}tr(M_t'M_t) \\ \frac{1}{N}\tilde{u}_t'\tilde{\bar{\bar{u}}}_t + \tilde{\bar{u}}_t'\tilde{\bar{u}}_t & \frac{-1}{N}\tilde{\bar{u}}_t'\tilde{\bar{\bar{u}}}_t & 0 \end{bmatrix}, \qquad g_t = \begin{bmatrix} \frac{1}{N}\tilde{u}_t'\tilde{u}_t \\ \frac{1}{N}\tilde{\bar{u}}_t'\tilde{\bar{u}}_t \\ \frac{1}{N}\tilde{u}_t'\tilde{\bar{u}}_t \end{bmatrix},$$

$t = 1, ..., T$. Here ν_t is the usual error associated with a sample of statistical realizations (i.e. it 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.[iii] That is, $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ that solve the non-linear optimization

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

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

$$\hat{\Phi}_t = (I_N - \hat{\rho}_t M_t)/\sqrt{\hat{\sigma}_t^2} \ .$$

(We could substitute $\hat{\Phi}_t$ for Φ_t and estimate equation (5), but we ultimately chose to restrict the model.) Let us called the $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ *unrestricted* estimates. We now consider feasible estimation of the fully restricted and partially restricted models discussed in the last section.

3.1 Feasible Estimation of the Fully Restricted System.

If we can assume that $M_1 = \ldots = M_T = M$, $\rho_1 = \ldots = \rho_T = \rho$, $\sigma_1^2 = \ldots = \sigma_T^2 = \sigma^2$ as before, then we can impose the assumption $M_1 = \ldots = M_T = M$ in equation (6) and estimate $\hat{\rho}_t$ and $\hat{\sigma}_t^2$, $t = 1, \ldots, T$ as above.[iv] Then average estimates of ρ and σ^2 are

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

We shall call these estimates the *fully restricted average* estimates. The estimates will be consistent as $N \to \infty$, so long as the restriction is true. These are essentially 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 sub-samples of the data of size N. These estimates imply

$$\hat{\Phi} = (I_N - \hat{\rho} M),$$

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

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 sub-samples, then the resulting parameter estimates should be more efficient than the average fully restricted estimates. One such estimate is based on the moment conditions:

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

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

(7) $\quad G[\rho, \rho^2, \sigma^2]' - g = \nu(\rho, \sigma^2),$

where,

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

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

$$(\tilde{\rho}, \tilde{\sigma}^2) = \arg\min_{r,s^2}[\nu(r,s^2)'\nu(r,s^2) : s^2 \geq 0]$$

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 the $\hat{\sigma}^2$ hinges on the fact that equation (7) exploits the information contained in *TN* observations and imposes a hypothetically correct restriction, while equation (6) exploits that contained in *N* observations over $t = 1, \ldots, T$, and no restriction. Again, $\tilde{\rho}$ and $\tilde{\sigma}^2$ imply $\tilde{\Phi}$, which can be inserted in equation (5); then fixed effect estimation of equation (5) with $\alpha^* = \iota_T \otimes \tilde{\Phi}\alpha$ is consistent for α^* and β. Estimation of α follows by transforming the estimate of α^* by $\tilde{\Phi}$.

3.2 Feasible Estimation of the Partially Restricted System.

For our 171 Indonesian rice farms observed over 6 periods, if we can assume that $M_1 = \ldots = M_6 = M$, $\rho_1 = \rho_3 = \rho_5 = \rho_W$, $\rho_2 = \rho_4 = \rho_6 = \rho_D$, $\sigma_1^2 = \sigma_3^2 = \sigma_5^2 = \sigma_W^2$, $\sigma_2^2 = \sigma_4^2 = \sigma_6^2 = \sigma_D^2$, then we can impose the assumption $M_1 = \ldots = M_6 = M$ 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*.

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

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

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

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

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

is consistent for α^* and β.

Define $\varepsilon_W' = [\varepsilon_1' \ \varepsilon_3' \ \varepsilon_5']$, $\varepsilon_D' = [\varepsilon_2' \ \varepsilon_4' \ \varepsilon_6']$, $\tilde{u}_W' = [\tilde{u}_1' \tilde{u}_3' \tilde{u}_5']$ and $\tilde{u}_D' = [\tilde{u}_2' \tilde{u}_4' \tilde{u}_6']$. Additionally, let $\tilde{\tilde{u}}_j = M\tilde{u}_j$, $\bar{\tilde{\tilde{u}}}_j = M\tilde{\bar{u}}_j$, $\bar{\varepsilon}_j = M\varepsilon_j$, and $\bar{\bar{\varepsilon}}_j = M\bar{\varepsilon}_j$, $j = W, D$. It follows analogously that the single stage estimates are:

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

where,

$$v_j(\rho_j, \sigma_j^2) = G_j[\rho_j, \rho_j^2, \sigma_j^2]' - g_j,\ j = W, D,$$

and where G_j and g_j are G_t and 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 $\tilde{\Phi}_j$ for wet and dry seasons, and fixed effect estimation of equation (5) is again consistent for α^* and β.

So to summarize, the unrestricted estimation procedure yields $\hat{\rho}_t$ and $\hat{\sigma}_t^2$. These estimates imply fully restricted average estimates ($\hat{\rho}$ and $\hat{\sigma}^2$) or partially restricted average estimates ($\hat{\rho}_j$ and $\hat{\sigma}_j^2$, $j = W, D$). These are two-stage estimates. Fully restricted moment estimates ($\tilde{\rho}$ and $\tilde{\sigma}^2$) and partially restricted moments estimates ($\tilde{\rho}_j$ and $\tilde{\sigma}_j^2$, $j = W, D$) are single stage estimates. We now illustrate these estimates in an empirical example.

4. Application to Indonesian Rice Farms

We now apply the estimators to 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, which has been extensively applied to this data set. Per Schmidt and Sickle a measure of technical efficiency for farm i based on this specification 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, M_t for the error process, which captures productivity spillovers across farms. The following section considers geographical and climate characteristics of the island of West Java, which motivate the construction of three different weighting matrixes used in subsequent analyses.

4.1 Geographical and Climatic Characteristics of West Java.

In 1977 the Indonesian Ministry of Agriculture began to survey 171 rice farms concerning farming practices over 6 (3 wet and 3 dry) growing seasons. The farms were selected from 6 villages located in the production area of Cimanuk River Basin in West Java. Of the six villages included in the sample, two are located on the north coast of the island in an area with average altitudes of 10-15 meters above sea level; these are classified as *lowland* villages. Another three villages are situated in a *highland* area (600-1100 meters) in the central part of West Java. The last village with average altitude of 375 meters is classified as a *midland* village (for lack of better terminology). The infrastructure in the Cimanuk River Basin is fairly heterogeneous. Some of the villages (in both high and lowland areas) lack reliable transportation systems with local roads being 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[vi].

Across the island of West Java average weather conditions are also highly variable. Annual temperatures average in the range of 72-84˚F (22-29˚C) with average humidity of 75%. However, the north coastal plains are usually hotter in the dry season, 94˚F (34˚C), and are more humid then the rest of the island. Highlands of the island receive an average annual rainfall of 156 inches (4000 mm) while lowlands of the north coast receive less than one forth of that amount, 35 inches (900 mm).

Based on these facts, we construct and perform our analysis using three different weighting matrixes $M1_t$, $M2_t$ and $M3_t$. The first one, $M1_t$, is based on the *inverse of geographical distance* between individual farms.[vii] 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 is unavailable and is therefore arbitrarily chosen to be 10 km.[viii] The *M1* 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.[ix] For $M2_t$ 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*. The last weighting scheme, $M3_t$, reflects the geographic and climate conditions of villages at different altitudes. The highland villages receive substantially more precipitation than low land villages and are characterized by lower average temperature than lowland villages. Moreover, certain villages are inaccessible during the rainy season. Therefore in the last weighting matrix, $M3_t$, we assign values based on differences in *altitudes* as follows.

m_{ijt} = 1.00 if farms i and j are in the same village

m_{ijt} = 0.75 if farms i and j are in different villages but at same altitudes.

m_{ijt} = 0.50 if farms i is highland and farm j is in midland.

m_{ijt} = 0.50 if farms i is midland and farm j is in lowland.

m_{ijt} = 0.00 if farms i is highland and farm j is in lowland.

For subsequent 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, all the weighting schemes are assumed time invariant, so the t subscript can be dropped.

4.2 Spatial Analysis of Indonesian Rice Farms.

We first estimate the standard fixed effect model of stochastic production frontier described by (1) and (2). 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; the omitted category represents that traditional varieties were planted. *DSS* equals 1 if it was a wet season, zero otherwise. Results are contained in column I of Table 1 and are based on the restriction that $\rho_1 = ... = \rho_6 = 0$. These results are identical to those contained 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 each of our three weighting schemes) exists in the data. As before, let the usual fixed effect residuals in period t be $\tilde{u}_t$. We employ two tests for spatial dependence, the first of which 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:

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

where N is the number of farms, S is the sum of all elements in weighting matrix M (S equals N if M is normalized so that sum of row elements equals one.) The null hypothesis for this test is "absence of spatial dependence".[x] 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 the asymptotic distribution for the Moran 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 Moran ϑ statistics 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:

$$E[\vartheta_t] = (N/S)tr(PM)/(N-k)$$

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

where P is the projection matrix $I_N - x(x'x)^{-1}x'$, and x is the demeaned exogenous variables from the standard model in equation (1). The test is conducted for each weighting scheme ($M1$, $M2$, $M3$) in each time period t = 1, ..., 6. The z_t -scores for weighting scheme $M1$ are contained in the top 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% confidence level, we reject the hypothesis of no spatial dependence based on weighting scheme *M1*. Test results for weighting schemes *M*2 and *M*3 were similar and are contained in the top rows (z_t) of Table 4 and Table 6, respectively.

As pointed out by an anonymous referee, 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, p 83). The test statistic:

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

is distributed χ_1^2 with critical values of 3.84 (95% level) and 6.63 (99% level). Results are contained in the last row of Tables 2, 4 and 6 for weighting schemes *M1*, *M2* and *M3*, respectively, and confirm the Moran *I* results: we reject the null in all cases.

Based on these test results, each of our proposed weighting schemes appears 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 contained in Tables 2, 4 and 6 for schemes *M1*, *M2* and *M3*, respectively. Note that for all 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 ($t = 2, 4, 6$). All autoregressive parameters are positive, and in only a single case does it exceed unity (scheme *M*3, period 3).[xi]

To identify parameter estimates for α^* and β in equation (5) we feasibly estimated the fully and partially-restricted systems described of sections 3.1 and 3.2, respectively. The fully restricted system, $\rho_1 = \ldots = \rho_6 = \rho$, is estimated using both the average autoregressive parameter, $\hat{\rho}$, and the moments autoregressive parameter, $\tilde{\rho}$, for each weighting scheme. Estimation results for $\hat{\rho}$ = 0.7248 and for $\tilde{\rho}$ = 1.0557 using weighting scheme *M1* are contained in Table 1, columns II and III, respectively. 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, given that ignoring the spatial dependence manifests itself as 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 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; this is currently under investigation by the authors. The results of the fully restricted model under weighting schemes *M2* and *M3* are contained in columns II and III in Tables 3 and 5, respectively. 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 wet and one for dry season. We calculate the *average parameter estimates* ($\hat{\rho}_W$, $\hat{\rho}_D$) and *the moments estimates* ($\tilde{\rho}_W$, $\tilde{\rho}_D$) for each weighting scheme. Fixed effect estimation results for ($\hat{\rho}_W$, $\hat{\rho}_D$) and for ($\tilde{\rho}_W$, $\tilde{\rho}_D$), based on weighting scheme *M1*, are contained in Table 1, columns IV and V, respectively. The differences between the average and moments parameter estimates are much less pronounced than the fully restricted case (compare estimates $\hat{\rho}_W$ = 0.7584 to $\tilde{\rho}_W$ = 0.8218, estimates $\hat{\rho}_D$ = 0.6914 to $\tilde{\rho}_D$ = 0.7476, and estimates $\hat{\rho}$ = 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 may suggest that the partially restricted model is 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). Of course the coefficient on the season variable (DSS) is not identified, since it is effectively time invariant now that the data set have been dichotomized into "wet" and "dry" sub-samples.[xii] 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 the 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 results are compelling and the coefficient *is* insignificant. Estimation results for weighting schemes M2 and M3 are similarly presented in columns IV and V of Table 3 and Table 5, respectively. Again, results are similar to scheme *M1* for this particular sample.

4.3 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 effect 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 was performed for the standard fixed effect model (corresponding to column I of Table 1) and for *the fully restricted moments estimator* (corresponding to column III of Tables 1, 3, and 5).[xiii] The idea was to see how the rankings differed between the standard model and the spatial model for each of the three weighting schemes. Order statistics for each model are contained in Table 7. 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 four farms with the highest technical efficiency, the four farms with the median technical efficiency, and the four farms with the lowest technical efficiency. Column 1 contains the farm number, column 2 contains the ordered estimates of farm-level technical efficiency, and column 3 contains the ordinal rankings for the standard fixed effect model (numbered 1 to 171). To see the effects of the spatial dependence on technical efficiency estimation, we also report the ordinal rankings for the same 12 farms for the fully restricted spatial model under weighting schemes *M1*, *M2* and *M3* in columns 4, 5 and 6, respectively.[xiv] 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 weighting schemes *M1* and *M2* and a rank of 4 under weighting scheme *M3*. 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. Most of 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 166, 166 and 108 for *M1*, *M2* and *M3*, respectively. These are potentially large changes in the median technical efficiency ranking, that could only be detected with a spatial analysis.

To further summarize the changes in the efficiency ranking in Table 7, we calculate *Spearman's rho* (r_s) for each weighting scheme, 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 7, $\delta_{164} = 0$ and $\delta_{15} = 86 - 62 = 24$. Here we always compare the rankings of the *M1*, *M2* and *M3* 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 contained in the last row of Table 7 and are on the order of 0.8 for each of the three weighting schemes. We can interpret this result as saying that only 80% of the rank statistic is preserved when we use a spatial weighting specification over the standard fixed effect specification.

To better understand the changes in technical efficiency under the various weighting schemes, we present some summary statistics and density estimates of estimates of the parameters, α_i. Technically, there is no distribution of α_i of which to speak, 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 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 make inferences on the effects of various weighting schemes on the estimates of technical efficiency is to make inferences on the estimates of α_i.

Table 8 contains summary statistics for the 171 estimates of α_i under various weighting schemes: standard fixed effect model, weighting scheme M1, weighting scheme M2, and weighting scheme M3. The results of Table 8 are best summarized in the kernel density plots for the various models contained in Figure 1. Density estimates are based on maximum likelihood cross validation bandwidth selection and a standard

Gaussian kernel.[xv] 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*, *M2* and *M3*). That is, without a spatial lag in the error, there is on average more technical inefficiency in the Indonesian rice farming industry than with a spatial lag. This is technical inefficiency in an absolute sense, since we are fixing α^* 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 may have suggested.

Of course in a relative sense, the technical inefficiency will be about the same across models, since in all models in Table 8 the difference between the maximal α_i and the average α_i is about the same (between 0.59 and 0.62). This is to say, that if we estimate maximal efficiency α^* as the maximum of the estimated α_i in each model, then the value of the difference between the estimate of maximal efficiency and the average of the estimates of α_i is about the same across models. However, as pointed out by Horrace and Schmidt (2000), this relative efficiency estimate will be biased in finite samples (small *N*), so perhaps viewing differences in relative technical inefficiency across models is illadvised in this data set, and our attention should be focused on the absolute differences across models, depicted in Figure 1.

5. Conclusions

This paper has presented a straight-forward generalization of the cross-sectional model of Kelejian and Prucha. However, the implications of the results transcend the spatial econometrics literature. Because economic agents and entities have finite lives, one cannot always rely on large T in economic panel data sets. There can be no denying that most panel data sets (with the exception of perhaps microeconomic financial data) have large N and small T. Moreover, the often high cost of empirical research necessarily impedes long, protracted data collection exercises; it is just cheaper to collect data sets with 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 $N\rightarrow\infty$. This is fine for estimating conditional means (the model's slope parameters). However, any second moment parameters (such as the elements of M) that embody cross-sectional dependence cannot be consistently estimated in the sense that they will necessarily rely on $T\rightarrow\infty$. This is unfortunate, because there can also be no denying that cross-sectional dependencies do exist in economic field data.

When faced with this dilemma, empirical economic researchers have two recourses, 1) collect more data or 2) 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 based on some geographic or economic proximity measures seem to be a reasonable and natural approach. The theoretical economic literature is rife with arguments for economic spill-overs, and spatial analysis provides a means to make these spill-overs 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.

We have presented two special cases of a more general model: the fully-restricted case and the partially-restricted case. Clearly, the possibilities for the partially restricted case are limitless. Our partial restriction based on wet and dry seasons was obviously data driven, so alternative restrictions for different types of data sets are possible. For example, infrastructure changes, catastrophic events or political dynamics may imply a different set of restrictions on the spatial correlation parameterization. Tests of these partial restrictions should be a high priority in subsequent research.

Notice that, dynamic spatial dependence in the second moment of our estimators has implications for dynamics in the first moment. The original model in equation (1) has a time invariant α vector, but the transformed model had time-varying α^*. 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 term 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 moment of α. For example, $\alpha_t = \mu + \delta t + \gamma t^2$ is a specification that has been considered. (For other examples see Cornwell, Schmidt and Sickles; Lee and Schmidt; Battese and Coelli; and Kumbhakar.) The models presented here create dynamic α through *second moment* conditions on u_t. What the implications of this difference will be for models of dynamic heterogeneity are unknown, but it is interesting to speculate. Clearly, the decision of whether to make α dynamic through the first moment or the second moment, should be driven by the application in mind. If it is safe to assume that errors are spherical, or if estimation efficiency is not a concern, or if there is strong evidence that dynamic effects are deterministic, then perhaps parameterizing the first moment is justified. However, if one believes that dynamic heterogeneity is caused by difference in spatial interactions, then perhaps parameterizing the second moment is appropriate. Either way, identification can only be achieved through a highly parametric specification, such as a parametric form for the dynamic α (in the first moment case) or a parametric form for the spatial weighting matrix (in the second moment case), so the decision may reduce to that of the selection of "the lesser of two evils", but this remains to be seen.

Additionally, spatial dependence may be a way to indirectly incorporate time-invariant regressors into a fixed effect model. For example, Horrace and Schmidt (1996) analyze the same data and incorporate five dummy variables for the six villages into a GLS or random effects specification, but they 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 second moment of the residual in the form of distances between villages in weighting matrix *M1* and in the form of contiguity or "common village" in matrix *M2*. While there are commonly employed techniques for incorporating time-invariant regressors into a fixed effect model, the research presented here provides analysts with an alternative means. (See Hausman and Taylor for a technique for time-invariant regressors.)

Our empirical example demonstrates the utility of this research; there are direct applications of this procedure to the stochastic frontier literature. We have illustrated the dangers associated with ignoring spatial dependence: a) a potential efficiency loss and b) the possibility of incorrect assessment of the technical efficiency rank statistics implied by the model. However, this modeling approach has much broader empirical implications. Any fixed effect model for a panel of data with fixed T could benefit from this type of analysis tool, if there are compelling reasons to believe that spatial dependence exists and if there are additional data (such as geographic or economic proximity measures) that motivate selection of a reasonable spatial weighting scheme.

Acknowledgements:
Research supported by the Office of the Vice Chancellor and the Center for Policy Research, Syracuse University. Thanks to Jeff Racine for the density estimates.

References

Aigner, D.J., C.A.K. Lovell And P. Schmidt. "Formulation and estimation of stochastic frontier production functions." *J. Econometrics* 6(July 1977): 21-37.

Anselin, L. *Spatial Econometrics: Methods and Models*. Boston: Klewer, 1988.

Anselin, L., A.K. Bera, R. Florax and M.J. Yoon. "A Simple Diagnostic Test for Spatial Dependence". *Reg. Sci. Urban Econ.* 26(February 1996): 77-104.

Battese and Coelli. "Frontier production functions, technical efficiency and panel data with application to paddy farmers in India." *J. Productiv. Anal.* 3(1992): 153-169.

Case, A.C. "Spatial patterns in household demand." *Econometrica* 59(July 1991): 953-966.

Cliff, A. and J. Ord. *Spatial Autocorrelation*. London: Pion, 1973

------. *Spatial Processes, Models and Applications*. London: Pion, 1981

------. " Testing for Spatial Autocorrelation Among Regression Residuals," *Geograph. Anal.* 4(1972): 267-84

Conley, T.G. "GMM Estimation with Cross-sectional Dependence." *J. 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." *J. Econometrics* 46(October-December 1990): 185-200.

Delong, J.B. and L.H. Summers. "Equipment investment and economic growth." *Quart. J. Econ.* 106(May 1991): 445-502.

Dubin, R.A. "Estimation of regression coefficients in the presence of spatially autocorrelated errors." *Rev. Econ. Statist.* 70(August 1988): 466-474

Erwidodo. "Panel Data Analysis on a Farm – Level Efficiency, Input Demand, Output Supply of Rice Farming in West Java, Indonesia. Dissertation. Dept. of Ag. Econ., Michigan State University, East Lansing, 1990

Fishback, P.V., W.C. Horrace, S.E. Kantor "Federal Programs in Times of Crisis: The Impact of the New Deal on Local Economies During the Great Depression." Working Paper, Dept. of Econ., University of Arizona, Tucson, 2002.

Geary , R. "The Contiguity Ratio and Statistical Mapping." *The Inc. Statistician* 5(1954):115-45.

Hausman, J.A. and W.E. Taylor. "Panel Data and Unobservable Individual Effects." *Econometrica* 61(November 1981):1377-98.

Horrace W.C. and P. Schmidt. "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models." *J. Productiv. Anal.* 7(July 1996): 257-282.

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

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

Kumbhakar, S.C. "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency." *J. of Econometrics*, 46(October-November 1990):201-11.

Lee, Y.H. and P. Schmidt. "A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency." *The Measurement of Productive Efficiency*. H.O. Fried et al., eds., pp. 3-67. New York: Oxford University Press, 1993.

Meeusen, W. And J. Van Den Broeck. "Efficient Estimation from Cobb-Douglas Production Functions with Composed Error." *Int. Econ. Rev.* 18(June 1977):435-444.

Moran, P. " The Interpretation of Statistical Maps." *J. Royal Stat. Soc.* B 10(1948):243-51.

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

Quah, D. "International Patterns of Growth: I. Persistence in Cross-County Disparities." Unpublished, London School of Economics, London, 1992.

Schmidt, P. and R.C. Sickles. "Production Frontiers and Panel Data." *J. Bus. Econ. Statist.* 2(October 1984):367-374.

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

Table 1. 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	0.1038	0.0998	0.1292	0.1248
	(0.030)	(0.025)	(0.024)	(0.024)	(0.024)
Urea	0.0918	0.0894	0.0901	0.1405	0.1440
	(0.021)	(0.018)	(0.017)	(0.015)	(0.015)
TSP	0.0892	0.0353	0.0244	0.0340	0.0307
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)
Labor	0.2431	0.2366	0.2379	0.2254	0.2204
	(0.032)	(0.029)	(0.028)	(0.026)	(0.026)
Land	0.4521	0.4879	0.4931	0.5046	0.5141
	(0.035)	(0.031)	(0.030)	(0.027)	(0.027)
DP	0.0338*	-0.0178*	-0.0298*	-0.0224*	-0.0212*
	(0.032)	(0.028)	(0.028)	(0.025)	(0.025)
DV1	0.1788	0.1084	0.0935	0.1250	0.1320
	(0.041)	(0.038)	(0.038)	(0.034)	(0.035)
DV2	0.1754	0.1060	0.0952	0.0917	0.0947
	(0.057)	(0.049)	(0.048)	(0.048)	(0.048)
DSS	0.0533	0.0759*	0.1062*	-	-
	(0.022)	(0.063)	(0.302)	-	-
R-sq	0.910228	0.9246	0.9271	0.9190	0.9177

Note: Numbers in parenthesis are standard errors. All estimates are significant at least at 5% significance level except those marked with an asterisk.

Table 2. Unrestricted Estimates, Weighting Scheme *M1*.

Time period	1	2	3	4	5	6
z_t	8.2420	6.0702	24.9468	26.4159	14.1925	12.3016
$\hat{\rho}_t$	0.6217	0.5217	0.8694	0.8375	0.7705	0.7101
$\hat{\sigma}_t^{\,2}$	0.0405	0.0828	0.0782	0.0713	0.0485	0.0658
LM_t	65.4080	30.5015	1461.0130	1680.6947	254.8534	175.9849

Table 3. 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-sq	0.910228	0.9240	0.9271	0.9171	0.9174

Note: Numbers in parenthesis are standard errors. All estimates are significant at least at 5% significance level except those marked with an asterisk.

Table 4. Unrestricted Estimates, Weighting Scheme *M2*.

Time period	1	2	3	4	5	6
z_t	7.4954	5.0873	23.5416	23.5057	13.6687	11.1836
$\hat{\rho}_t$	0.5682	0.4803	0.7875	0.7889	0.6875	0.6501
$\hat{\sigma}_t^{\ 2}$	0.0407	0.0842	0.0774	0.0718	0.0481	0.0661
LM_t	57.0245	21.6164	1409.1248	1386.3032	256.7318	153.0266

Table 5. Weighting Scheme *M3* – Altitude Differences

	Standard FE Model I	Fully Restricted Average II	Fully Restricted Moment III	Partially Restricted Average IV	Partially Restricted Moment V
$\hat{\rho}$	-	0.8648	-	-	-
$\tilde{\rho}$	-	-	1.0673	-	-
$\hat{\rho}_w$	-	-	-	0.9130	-
$\hat{\rho}_D$	-	-	-	0.8166	-
$\tilde{\rho}_W$	-	-	-	-	0.9929
$\tilde{\rho}_D$	-	-	-	-	0.8446
Seed	0.1208	0.1114	0.1095	0.1430	0.1500
	(0.030)	(0.025)	(0.025)	(0.023)	(0.023)
Urea	0.0918	0.0943	0.0963	0.1377	0.1342
	(0.021)	(0.018)	(0.018)	(0.015)	(0.015)
TSP	0.0892	0.0325	0.0283	0.0330	0.0341
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)
Labor	0.2431	0.2489	0.2522	0.2362	0.2407
	(0.032)	(0.028)	(0.028)	(0.026)	(0.026)
Land	0.4521	0.4794	0.4800	0.4902	0.4803
	(0.035)	(0.031)	(0.030)	(0.028)	(0.027)
DP	0.0338*	-0.0185*	-0.0256*	0.0033*	-0.0057*
	(0.032)	(0.028)	(0.028)	(0.023)	(0.023)
DV1	0.1788	0.1216	0.1203	0.1631	0.1610
	(0.041)	(0.037)	(0.036)	(0.034)	(0.035)
DV2	0.1754	0.1304	0.1309	0.1205	0.1168
	(0.057)	(0.049)	(0.048)	(0.050)	(0.048)
DSS	0.0533	0.0804*	0.0683*	-	-
	(0.022)	(0.128)	(0.254)	-	-
R-sq	0.910228	0.9280	0.9289	0.9180	0.9196

Note: Numbers in parenthesis are standard errors. All estimates are significant at least at 5% significance level except those marked with an asterisk.

Table 6. Unrestricted Estimates, Weighting Scheme *M3*.

Time period	1	2	3	4	5	6
z_t	14.0899	15.9659	57.4523	54.5444	25.3519	26.3768
$\hat{\rho}_t$	0.7510	0.7129	1.0384	0.9044	0.9497	0.8325
$\hat{\sigma}_t^2$	0.0415	0.0811	0.0765	0.0764	0.0501	0.0661
LM_t	89.4159	119.9479	3658.9348	3222.0965	368.5170	413.4162

Table 7. Technical Efficiency Orders Statistics, Various Models

Standard FE Model			**Spatial Models**		
Farm #	Standard FE Efficiency	Standard FE Model	Weight Scheme *M1*	Weight Scheme *M2*	Weight Scheme *M3*
164	100%	1	1	1	1
118	93.23%	2	2	2	3
163	93.03%	3	3	3	2
152	89.93%	4	6	6	4
13	55.62%	84	106	106	114
166	55.47%	85	116	116	108
15	55.40%	86	62	62	72
40	55.35%	87	54	54	64
86	39.80%	168	165	165	166
143	38.37%	169	169	169	170
117	37.90%	170	168	168	168
45	36.55%	171	171	171	171
	r_s:	1.0000	0.8027	0.8095	0.8674

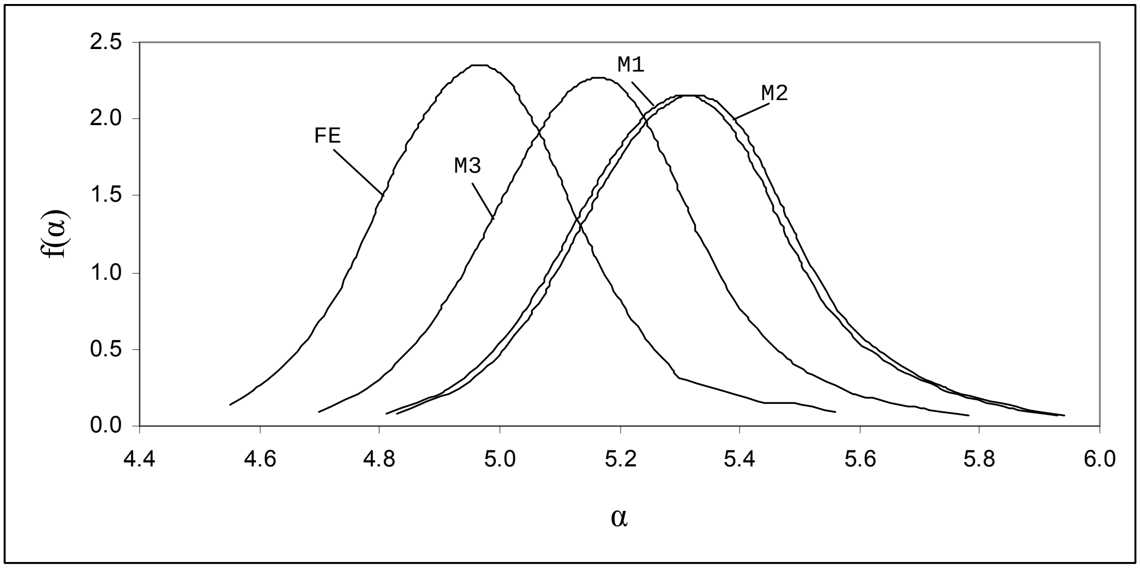
Spatial results are for the fully restricted moments estimator.
Technical Efficiency for farm $i = \exp\{\alpha_i - \max_j \alpha_j\}$.

Table 8. Summary Statistics for α_i

		-------Spatial Models-------		
Statistic	Standard FE Model	Weight Scheme *M1*	Weight Scheme *M2*	Weight Scheme *M3*
Average α_i	4.97	5.31	5.33	5.17
Median α_i	4.97	5.30	5.32	5.16
Std. Dev. α_i	0.18	0.19	0.19	0.18
Max α_i	5.56	5.93	5.94	5.78
Min α_i	4.55	4.81	4.83	4.70

Spatial results are for the fully restricted moments estimator.

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, M3 = M3 Weighting Scheme

Appendix
Assumptions 3, 4, and 5 from Kelejian and Prucha (1999). Let $P(\rho_t) = (I_N - \rho_t M_t)^{-1}$ 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 $\tilde{u}_t$. There exists finite dimensional random vectors d_{it} and Δ_t such that $|\tilde{u}_{it} - u_{it}| \leq ||d_{it}||\ ||\Delta_t||$ with $N^{-1}\Sigma_i||d_{it}||^{2+\delta} = O_p(1)$ for some $\delta > 0$ and $N^{-1/2}\Sigma_i||\Delta_t|| = O_p(1)$.

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

End Notes

---

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

[ii] We present no proofs of our results, because they are all straight-forward extensions of Kelejian and Prucha's proofs. Proofs are available from the second author upon request.

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

[iv] 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 boot-strapped. In the sequel we use the Moran I test and the Lagrange Multiplier test to test the significance of the over-all weighting scheme in each period.

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

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

[vii] 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.

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

[ix] Both Moran and Geary advanced initial measures of spatial dependence (spatial correlation) that were based on the notion of binary contiguity between spatial units. Underlying spatial structure is expressed by 0-1 values: if spatial units have common border (are contiguous) value of 1 is assigned.

[x] 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 absence of spatial dependence, a precise expression for the alternative does not exist.

[xi] Although there is no widely accepted interpretation of this phenomenon, we view an estimated ρ greater than unity as a specification error. Therefore, weighting scheme *M3* may not be justified.

[xii] Note that even though the time dimension have 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".

[xiii] We did not consider comparing the rankings of the partially-restricted system, since the α estimates are based on only 3 observations and in the other cases (standard and fully restricted systems) they are based on 6 observations.

[xiv] To save space we do not report the actual technical efficiency estimates for the spatial models, only the efficiency ranking.

[xv] Details of the kernel estimations are available by request from the second author.