



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

**GENERAL MOMENT AND QUASI-
MAXIMUM LIKELIHOOD ESTIMATION
OF A SPATIALLY AUTOCORRELATED
SYSTEM OF EQUATIONS: AN EMPIRICAL
EXAMPLE USING ON-FARM PRECISION
AGRICULTURE DATA**

by
Dayton M. Lambert, Gary L. Malzer,
and James Lowenberg-DeBoer

Staff Paper #04-02

February 2004

Dept. of Agricultural Economics

Purdue University

Purdue University is committed to the policy that all persons shall have equal access to its programs and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

GENERAL MOMENT AND QUASI-MAXIMUM LIKELIHOOD ESTIMATION OF A SPATIALLY AUTOCORRELATED SYSTEM OF EQUATIONS: AN EMPIRICAL EXAMPLE USING ON-FARM PRECISION AGRICULTURE DATA

by

Dayton M. Lambert^a

Gary Malzer^b

and

James Lowenberg-DeBoer^a

a Department of Agricultural Economics, Purdue University
West Lafayette, Indiana, 47907-1145

b Department of Soil, Water, and Climate, University of Minnesota
Rm. S401 Soil Science Building
1991 Upper Buford Circle, St. Paul MN 55108

lamberd@purdue.edu

lowenbej@purdue.edu

malze001@tc.umn.edu

Staff Paper #04-02

February 2004

Abstract

This paper compares two estimation methods of a three-stage least squares (3SLS) system of equations, corrected for spatial autocorrelation. The modeling approach is novel in that it is an extension of Anselin's (1988) seemingly unrelated regression (SUR) space-time spatial error model for panel data. An empirical comparison of the quasi-maximum likelihood (QML) estimation of the equation system, and Kelejian and Prucha's general moments (GM) estimation approach is presented. The model and estimation procedures introduced in this study are easily extended to other economic, agronomic, or biological models that must incorporate spatial and temporal effects in the model specification, and overcome simultaneous equation bias. The empirical example used in this study falls in the realm of production economics: on-farm production data is used to optimize input rates across time and space. This data is the product of on-farm, site-specific manure management research at the University of Minnesota.

Keywords: General moment, Quasi-maximum likelihood, three stage least squares, spatial regression, on-farm data.

Copyright © 2004 by Dayton M. Lambert, Gary L. Malzer, and James Lowenberg-DeBoer. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

**GENERAL MOMENT AND QUASI-MAXIMUM LIKELIHOOD
ESTIMATION OF A SPATIALLY AUTOCORRELATED SYSTEM
OF EQUATIONS: AN EMPIRICAL EXAMPLE USING ON-FARM
PRECISION
AGRICULTURE DATA**

by
Dayton M. Lambert
Gary Malzer,
and
James Lowenberg-DeBoer

Introduction

Recently, much attention has been paid towards modeling spatial effects in applied economics. The estimation problems caused by spatial dependence have been well known for some time: assumptions about the independence of observations are often violated when data is spatial, much as they are in time series contexts. Along with the newly received understanding of, and sensitivity about the effects of spatial processes by economists, detection of and correction for spatial dependence in economic data is no longer the exception in applied economics. Geo-referenced data is now widely available for research purposes, and faster computers supporting data-visualizing GIS software provide researchers with new perspectives on data generating processes (DGP) by inclusion of spatial dimensions. For the applied economist, modeling spatial processes is also intuitively appealing since data used by economists is the by-product of micro-transactions between agents occupying definite locations in a continuum of time. Macro-level data, such as global trade patterns, national income, financial transactions, unemployment spells, and the like are no exception, since realization of these aggregate-level data are driven by a nexus of utility-maximizing agents and profit-maximizing firms at micro-levels.

The econometric time series literature is rich and detailed, but the breadth and depth of applied spatial econometric literature available to econometricians is not so profound. Yet, since the last decade, spatial regression methods have been adapted to fit applied economic and agro-economic research paradigms. The spatial regression tools currently available to economists were originally developed in the ecological sciences, urban and regional geography, agronomy, and the geological sciences. Needless to say, the methods developed in these disciplines were not designed for the kinds of regression models applied economists commonly employ (for example, estimation of demand systems, trade models and systems of supply-demand equations, panel data, discrete choice and proportional hazard models, and vector auto-regression (VAR), or other time-series related models). Accordingly, the application of spatial analysis developed in these other disciplines has been modified to varying degrees to fit economic theory and econometric methods, depending on the applied problem at hand.

This paper proceeds as follows. The literature review summarizes some recent applications of spatial econometrics in applied economics, agricultural economics, and related fields. The problems associated with the inclusion of time-effects into spatial models (or vice-

versa, as in the case of VAR models) are briefly discussed, along with some proposed solutions towards identification and estimation of space-time models. Next, two estimation methods are discussed in detail with which to estimate a system of equations in general, and 3SLS in particular – a quasi-maximum likelihood (QML, Anselin, 1988) and a general moments (GM, Kelejian and Prucha, 1999, 2004) approach. The section following describes the data and the model used to compare the results produced by the QML and GM estimation procedures. The production economic data used in the example are the results of a two-year on-farm, variable rate input trial. Results are presented, followed by a concluding section.

Literature Review

There are many approaches towards modeling spatial processes. Anselin's (1988) *Spatial Econometrics* and Cressie's (1993) *Statistics for Spatial Data* are arguably the most exhaustive reviews of the asymptotic properties of estimators, estimation procedures, and examples of spatial regression methods at work. Although many of the spatial models described by Cressie are easily modified into useful econometric models, Anselin's review has direct applications in economics.

In general, most of the recent economic studies applying spatial methods follow one of two models extensively detailed by Anselin (1988, 2002)¹ – spatial *lag* or spatial *error* models. For spatial lag processes, the familiar regression model becomes $\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$; with ρ as the autoregressive moving average parameter for neighboring y_j 's. The spatial autoregressive (SAR) error model is specified as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ with $\boldsymbol{\varepsilon} = \lambda \mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u}$, where \mathbf{u} represents well-behaved, non-heteroskedastic, uncorrelated errors. \mathbf{W} is a positive definite, $n \times n$ matrix of spatial weights. Spatial structure is necessarily imposed on data. That is, the researcher chooses the elements of \mathbf{W} . For example, \mathbf{W} may be a matrix of economic distances (Conley, 1999; Conley and Topa, 2002; Aten, 1997), zeros and ones identifying neighborhood contiguity (Anselin, Bongiovanni, and Lowenberg-DeBoer, 2004), inverse or Euclidean distances, or the proportion of shared boundaries between districts, counties, or states (Ord 1975). The required properties of \mathbf{W} are that it is invertible, and each of its eigenvalues is positive and less than unity. Details about the properties of \mathbf{W} are found in Anselin (1988) and Bell and Bockstael (2000).

In the agricultural economic literature, one of the earlier applications of the single equation SAR model is found in Benirschka and Binkley (1994), where they modeled land price variations as a function of distance to markets. More recently, Roe, Irwin, and Sharp (2002) modeled the influence of spatial relations on economic agglomeration in the hog industry using a spatial error model. A natural extension of spatial econometrics is evaluation of on-farm experiments. Florax, Voortman, and Brouwer (2002) used a spatial lag model to capture local variations in soil characteristics, and their impact on millet yield in Niger. Anselin, Bongiovanni, and Lowenberg-DeBoer (2004) used a SAR model to evaluate yield monitor data for an on-farm variable rate nitrogen trial (VRT-N) in Argentina. Heermann et al. (2002) also used the SAR specification in their evaluation of precision irrigation systems on corn response in Colorado.

¹ Recently, Conley (1999) introduced a non-parametric general method of moments (GMM) approach towards estimating spatial effects in economic data. But for exposition, this study focuses exclusively on estimation of SAR error models.

The common thread between these studies is the cross-sectional attributes of the data. The usual conceptual and computational problems associated with estimating spatial models are compounded with multiple periods and panel series data. As a conceptual example, in time series applications, the DGP is assumed to be unidirectional over time. On the other hand, spatial processes are multidirectional.

In general, spatial processes are assumed to follow a Markov random field. Spatial effects might be modeled two dimensionally (like a “checkerboard”) in the case of the epidemiologist studying disease incidence between counties, or three-dimensionally in the case of the oil or mineral wildcatter. Adding the temporal dimension to spatial process requires one to imagine that a given tessellation (which is generally fixed over space) moves through time. The data observed by the economist (for example, transactions between individuals) is generated along these points. Not only are observational units connected to other units on a network; they are connected to themselves through period linkages. Immediate examples include interpretation of yield monitor data over several crop cycles; market transactions between economic agents; watersheds, land tenure, and other ecosystem-related studies; changes in biological populations over space and time; and international trade flows. Computationally, the problem is compounded by (possible) inclusion of extra parameters that must be estimated (as in the case of Pace et al., 2000), an increase in the size of spatial weighting matrices (Lence and Mishra, 2003; Druska and Horrace, 2004; Hardie et al., 2001), or linkage of spatial and temporal effects through a system of equations or temporal effects through random error components (Anselin, 1988).

Recently, these conceptual and computational challenges have been addressed in several applied economic studies. In general, there are four approaches available to applied economists to concomitantly model spatial and temporal effects; (i) error component models (ECM; Elhorst, 2001; Baltagi, Song, and Koh, 2003), especially for panel data and discrete choice spatial models (Munroe, Southworth, and Tucker, 2002; Swinton, 2002; Pinske and Slade, 1998; Holloway, Shankar, and Rahman, 2002), (ii) seemingly unrelated regression (SUR, Anselin, 1988), (iii) direct estimation of spatial and temporal autoregressive (AR) parameters in model specifications, and (iv) inclusion of spatial weights matrices vis-à-vis parameter restrictions in VAR models (LeSage and Pan, 1995; Chen and Conley, 2001; Leonard and Somy, 1997; Kamarianakis, 2003; Giacomini and Granger, 2004).

Empirical applications using a system of equations or SUR approach are less frequent in the literature. Benirschka and Binkley (1995) used an instrumental variable approach to model space-time effects in their study of optimal storage and marketing of grain over space and time. However, cross-equation covariance was not explicitly modeled in that study. The SUR approach taken in this study follows Anselin’s program: each equation corresponds to a specific time period. This allows spatial processes to be different in every period: it is assumed that temporal effects interacting over a fixed spatial arrangement may have different effects every period. The asymptotic properties of the QML estimators of the spatial SUR are given in Anselin (1988), along with an empirical example. Kelejian and Prucha (2004) established the asymptotic properties of GM estimators for systems of equations, but there are no empirical examples in the economic literature comparing SUR-QML and GM estimators. Additionally, the GM-systems approach in Kelejian and Prucha only considers cross-section systems of equations. Their model is extended here to include equations representing time blocks in this study. Like the SUR

approach in Anselin, and the Kelejian-Prucha GM method, a spatial AR term (λ) is estimated for each equation.

Pace et al. (2000), Hardie et al. (2002), Silveira-Neto and Azzoni (2003), and Lence and Mishra (2003) provide recent examples of inclusion of spatial AR terms into conventional panel data models. However, unlike the system approach taken here, these studies restricted λ to be identical across all periods. Druska and Horrace (2004) used Kelejian and Prucha's GM procedure to estimate a panel data series, but they included a spatial AR term for each period. However, they did not link inter-temporal (t) effects through the AR terms vis-à-vis a cross-equation covariance matrix. With the SUR approach, inter-period linkage is achieved through the cross-equation covariance matrix as a function of λ_t 's, $\Sigma(\Lambda)$, where Λ is a $T \times T$ block diagonal matrix of spatial AR terms for each time period equation (Anselin, 1988, page 143).

The remainder of this paper builds specifically upon the Cliff-Ord-type, single equation SAR model, and its modification into a system of equations for estimation of space-time interactions with endogenous and instrumental variables. Since the model includes endogenous variables on the right hand side of the estimating equations, 3SLS must be used in estimating the system. Two estimation methods already available to researchers are further developed and compared: the quasi-maximum likelihood (QML) approach (Anselin, 1988), and Kelejian and Prucha's GM procedure. The GM procedure presented by Kelejian and Prucha (1999, 2004) is modified by iterating over the cross-equation covariance matrix, Σ . Likewise; this iterative approach is used in QML computations.

Specification of a Linear System Space-Time System of Equations Model

This section specifies the system of equations used to estimate profitability of variable rate input management in Minnesota. As is well known in the econometric literature, single-equation estimators are not asymptotically efficient if the structural errors are correlated across equations. Kelejian and Prucha (2004) consider a general spatial model that includes spatial lag and error processes. In their approach, the normality assumption required by QML estimation is relaxed. Like the spatial SUR described in Anselin (1988, page 143), the model developed below only considers spatial error processes.

Quasi-maximum likelihood estimation of the 3SLS space-time model

Consider extending the Cliff-Ord single equation SAR model as the following simultaneous system of $p = 1, \dots, P$ equations, each with n observations, occurring in $t = 1, \dots, T$ cross-sectional units (or periods):

$$(1) \quad \begin{aligned} \mathbf{Y}_{pt} &= \mathbf{Y}_{(pt)}\gamma_{pt} + \mathbf{X}_{(pt)}\beta_{pt} + \mathbf{e}_{pt}, \mathbf{e}_{pt} = \lambda_{pt}\mathbf{W}_n\mathbf{e}_{pt} + \mathbf{u}_{pt} \\ &= \mathbf{M}_{(pt)}\delta_{pt} + \mathbf{e}_{pt}, \mathbf{e}_{pt} = \lambda_{pt}\mathbf{W}_n\mathbf{e}_{pt} + \mathbf{u}_{pt} \end{aligned}$$

In vector notation

$$(2) \quad \mathbf{Y}_v = \mathbf{M}\delta + \mathbf{e}_v, \mathbf{e}_v = \lambda_v\mathbf{W}_n\mathbf{e}_v + \mathbf{u}_v$$

$$(3) \quad E[\mathbf{u}_{pt} \mathbf{u}_{qs}'] = \sigma_{pt,qs} \mathbf{I}_n$$

where $\mathbf{Y}_v = (\mathbf{Y}'_{11}, \mathbf{Y}'_{21}, \dots, \mathbf{Y}'_{PT})'$, $\boldsymbol{\delta} = (\boldsymbol{\delta}'_{11}, \boldsymbol{\delta}'_{21}, \dots, \boldsymbol{\delta}'_{PT})'$, and $\mathbf{e}_v = (\mathbf{e}'_{11}, \mathbf{e}'_{21}, \dots, \mathbf{e}'_{PT})'$, and the block diagonal matrix \mathbf{M} , wherein each $\mathbf{M}_{(pt)}$ may contain one or more endogenous variables (Mittelhammer, Judge, and Miller, 2000). In this system, the errors follow a spatial autoregressive process in each equation, and are correlated across equations and time blocks. Each of the P equations has an AR term in each period t , λ_t . Re-writing the \mathbf{e}_{pt} errors as $\mathbf{e}_{pt} = (\mathbf{I}_n - \lambda_{pt} \mathbf{W}_n)^{-1} \mathbf{u}_{pt}$, it follows that

$$(4) \quad E[\mathbf{e}_{pt} \mathbf{e}_{qs}'] = \sigma_{pt,qs} \left[\mathbf{G}_{pt}' \mathbf{G}_{pt} \right]^{-1} : \mathbf{G}_{pt} = (\mathbf{I}_n - \lambda_{pt} \mathbf{W}_n)$$

For the full system model, the error covariance matrix ($\boldsymbol{\Omega}$) can be expressed as $\boldsymbol{\Omega} = \mathbf{G}^{-1} (\boldsymbol{\Sigma} \otimes \mathbf{I}_n) \mathbf{G}'^{-1}$, where \mathbf{G} is an $nPT \times nPT$ block diagonal matrix. In simpler notation, $\mathbf{G} = (\mathbf{I}_{nPT} - \boldsymbol{\Lambda} \otimes \mathbf{W}_n)$, where $\boldsymbol{\Lambda}$ is a $PT \times PT$ block diagonal matrix of AR terms for the p -th equation in the t -th period.

QML estimators derived under the normal approximation may have all the desired statistical properties (for example, consistency and asymptotic normality), even if the approximation is incorrect (Gourieroux, Monfort, and Trognon, 1984). Under quasi-likelihood theory, the normal likelihood function is used as an approximation to the true (but unknown) likelihood function when the errors may be distinctively non-normal, even in very large sample sizes (Mittelhammer, Judge, and Miller, 2000). The information in the data and model can be summarized by the following log likelihood objective function (following Anselin, 1988, page 143; Mittelhammer, Judge, and Miller, 2000, page 465):

$$(5) \quad \ln \ell(\boldsymbol{\delta}, \boldsymbol{\Lambda}, \boldsymbol{\Sigma}; \mathbf{Y}_v, \mathbf{X}) = -\left(\frac{n}{2}\right) \ln |\boldsymbol{\Sigma}| + \sum_{t=1}^T \sum_{p=1}^P \ln |\mathbf{I}_n - \lambda_{pt} \mathbf{W}_n| \\ - \frac{1}{2} (\mathbf{Y}_v - \mathbf{M}\boldsymbol{\delta})' \mathbf{G}' [\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_n] \mathbf{G} (\mathbf{Y}_v - \mathbf{M}\boldsymbol{\delta})$$

The error variance is a function of λ and σ (the upper triangular elements of the $PT \times PT$ covariance matrix, $\boldsymbol{\Sigma}$). That is, $\boldsymbol{\Sigma} \equiv \boldsymbol{\Sigma}(\boldsymbol{\Lambda}) = n^{-1} \mathbf{Z}' \mathbf{Z}$, where \mathbf{Z} is a PT vector of transformed residuals; $\mathbf{Z} = [\mathbf{z}_{11}, \mathbf{z}_{12}, \dots, \mathbf{z}_{PT}]$, $\mathbf{z}_{pt} = (\mathbf{I}_n - \lambda_{pt} \mathbf{W}_n) \mathbf{u}_{pt}$; and $\mathbf{u}_{pt} = \mathbf{y}_{pt} - \mathbf{y}_{pt} \boldsymbol{\gamma}_{pt} - \mathbf{x}_{pt} \boldsymbol{\beta}_{pt}$.

The estimates of $\boldsymbol{\Lambda}$ are found as a solution to the system of PT nonlinear equations:

$$(6) \quad tr(\mathbf{W}_n (\mathbf{I}_n - \lambda_{pt} \mathbf{W}_n)^{-1}) - \left[\sum_{q=1}^P \sum_{s=1}^T \sigma_{ps}^{-1} [\mathbf{u}_{pt}' \mathbf{W}_n' (\mathbf{I}_n - \lambda_{ps} \mathbf{W}_n) \mathbf{u}_{qs}] \right] = 0$$

where $\sigma_{p,t}$ is the p, t -th element of $\boldsymbol{\Sigma}(\boldsymbol{\Lambda})$. This relation between the necessary conditions for λ_{pt} demonstrates that estimation of one spatial AR term in the p -th equation and t -th period is conditional upon the values of the other AR terms. Temporal and spatial autocorrelation between equations and periods is thereby linked through the matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\Sigma}$. The steps for maximizing

(5) and solving the system of non-linear equations in (6) follow Anselin (1988, page 183, 184). Noting that $\hat{\mathbf{M}} = (\mathbf{I}_n \otimes \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{M}$ (Mittelhammer, Judge, and Miller, 2000), the 3SLS estimator that solves (5) is:

$$(7) \quad \hat{\delta} = \left\{ \hat{\mathbf{M}}' \mathbf{G}'(\hat{\Lambda}) \left[\hat{\Sigma}(\hat{\Lambda})^{-1} \otimes \mathbf{I}_n \right] \mathbf{G}(\hat{\Lambda}) \hat{\mathbf{M}} \right\}^{-1} \hat{\mathbf{M}}' \mathbf{G}'(\hat{\Lambda}) \left[\hat{\Sigma}(\hat{\Lambda})^{-1} \otimes \mathbf{I}_n \right] \mathbf{G}(\hat{\Lambda}) \mathbf{Y}_v$$

When the λ_{pt} 's are recovered from equation 6, along with the final estimate of $\Sigma(\Lambda)$, $\hat{\delta}$ is estimated. At this point, the information matrix can be constructed (Anselin, 1988, page 161), and standard errors of the parameters are estimated.

General Moments (GM) Estimation of the 3SLS model

Maximum likelihood estimation of the spatial autoregression error model may not always be computationally feasible with larger spatially referenced data sets, systems of equations, or regression models incorporating time and space. The main obstacle presents itself during maximization of the log likelihood function where the determinant of the $n \times n$ matrix $\ln|\mathbf{I}_n - \lambda \mathbf{W}_n|$ must be calculated in every iteration. Ord (1975) showed that computational time could be reduced by rewriting this term as $\ln \sum_{i=1}^n \text{abs}(1 - \lambda \omega_i)$, where ω_i is the i -th eigenvalue of \mathbf{W}_n . Since the eigenvalues of \mathbf{W}_n only have to be found once, computation time is reduced for QML estimation. However, even in larger samples, re-writing the problem this way may not decrease estimation time, and estimates of the eigenvalues may not be accurate (for example, see Bell and Bockstael, 2000). This may be especially true with larger panel data sets, or data modeled using systems of equations.

Kelejian and Prucha (1999) proposed an alternative method for estimating the SAR model that does not require estimation of eigenvalues of \mathbf{W}_n or the log determinant, $\ln|\mathbf{I}_n - \lambda \mathbf{W}_n|$. Recent empirical applications using the GM approach towards estimation of SAR models include Druska and Horrace (2004), and Bell and Bockstael (2000). Using the above notation, the following system of general moment equations solves for λ_{pt} for a single equation in period t :

$$(8) \quad \Gamma_{pt} = \frac{1}{n} \begin{bmatrix} 2\mathbf{u}'_{pt} \mathbf{W}_n \mathbf{u}_{pt} & -(\mathbf{W}_n \mathbf{u}_{pt})' \mathbf{W}_n \mathbf{u}_{pt} & n \\ 2(\mathbf{W}_n^2 \mathbf{u}_{pt})' \mathbf{W}_n \mathbf{u}_{pt} & -(\mathbf{W}_n^2 \mathbf{u}_{pt})' \mathbf{W}_n^2 \mathbf{u}_{pt} & \text{tr}(\mathbf{W}_n^2) \\ \left(\mathbf{u}'_{pt} \mathbf{W}_n^2 \mathbf{u}_{pt} + [\mathbf{W}_n \mathbf{u}_{pt}]' \mathbf{W}_n \mathbf{u}_{pt} \right) & -(\mathbf{W}_n \mathbf{u}_{pt})' \mathbf{W}_n^2 \mathbf{u}_{pt} & 0 \end{bmatrix}$$

$$(9) \quad \Xi_{pt} = [\lambda_{pt}, \lambda_{pt}^2, \sigma_{pt}^2]'$$

$$(10) \quad \gamma_{pt} = \frac{1}{n} \left[\mathbf{u}'_{pt} \mathbf{u}_{pt}, (\mathbf{W}_n \mathbf{u}_{pt})' \mathbf{W}_n \mathbf{u}_{pt}, \mathbf{u}'_{pt} \mathbf{W}_n \mathbf{u}_{pt} \right]'$$

$$(11) \quad \arg \min_{\theta_{pt}} \left\{ \left(\Gamma_{pt} \Xi_{pt} - \gamma_{pt} \right)' \left(\Gamma_{pt} \Xi_{pt} - \gamma_{pt} \right) : \lambda_{pt} \in [-1, 1], \sigma_{pt}^2 \geq 0 \right\}$$

One approach is to use a two-step estimation procedure: the elements in Λ are estimated for each equation and every time block. Once λ_{pt} is recovered, $\Sigma(\Lambda)$ and then δ can be estimated using the identity in (7). An alternative approach would entail an iterative procedure, taking similar steps outlined to estimate the parameters in (5) (see Anselin. 1988, page 144). For example:

- (i) Two-stage least squares estimates of \mathbf{u}_{pt} are obtained from (2);
- (ii) Ξ_{pt} 's are estimated solving the system of moment equations, and a new set of residuals (\mathbf{u}_{pt}^*) are constructed using (2) and (7);
- (iii) Σ is re-estimated as $\Sigma(\Lambda)$;
- (iv) δ 's are recovered using equation 7, conditional upon $\Sigma(\Lambda)$;
- (v) A new set of residuals is estimated;
- (vi) Steps (ii) – (v) are repeated until a specified convergence criterion is achieved.

Once the convergence criterion is obtained, the information matrix for the system is constructed, conditional upon $\hat{\Lambda}$, and standard errors of the estimates are recovered.

An Empirical Example

The purpose of this section is to use the estimators described in section 3 to estimate actual site-specific crop response functions (SSCRF, Bullock, Lowenberg-DeBoer, and Swinton, 2002) using real spatial-temporal observations that are representative of the kind of data commonly used by agricultural researchers. An on-farm trial was established near Sleepy Eye, Minnesota to determine if variable rate manure (VRM) could be used to increase agronomic profit and reduce the potential for non-point source environmental problems. The objectives of this trial are to compare the returns to VRM with returns from a whole-field manure management strategy. Each manure management strategy is complemented by a one of two soil management strategies: do nothing, or use soil test information to adjust K, P, or lime site-specifically. In all, there are four scenarios compared.

The variable rate manure experiment was conducted in cooperation with Christensen farms, near Sleepy Eye MN. Corn grown during the 1999 season was followed by soybean. Four rates of liquid swine manure, including a check strip (0, 2000, 4000, 6000, and 8000 gal acre⁻¹) were applied over a 10-acre field in constant rate strips. Manure was only applied before the corn-growing season. No manure was applied prior to planting soybean. Manure was applied via surface broadcast, then immediately incorporated with double discs attached to the applicator. Yield data was collected in 15-m segments for corn and soybean crops. Grain yield was measured from the center row of each treatment strip using a Massey Ferguson plot combine equipped with a ground distance monitor and computerized Harvest Master weigh-all (Harvest Master, Logon UT). Every 15-m, the combine was stopped and the grain weighed. For more details, see Lambert, Malzer, and Lowenberg-DeBoer, 2003.

Estimating Equations for Crop Production Function

A second-degree polynomial response function (Dillon and Anderson, 1990) is assumed to describe corn and soybean response to manure. This particular functional has the advantage in that it can be interpreted as second order approximations of any crop response function. For crop production analyses, quadratic functions are useful since they are concave functions with $f' > 0$ and $f'' < 0$. This allows diminishing marginal returns to inputs applied, and allows for the possibility that crop growth can be thwarted by input applications above biophysically optimal levels. The estimation of economically optimal input rates (EOR's) is also possible since a closed-form solution to f' exists.

The notation used to describe the system of corn and soybean production functions is listed in Table 1. The following equations model site-specific crop response of corn and soybean to manure as a system of production functions:

$$(12) \quad \mathbf{y} = \mathbf{XA} + \hat{\mathbf{C}}\boldsymbol{\gamma} + \mathbf{e}, \mathbf{e} = \lambda_{at}\mathbf{We} + \mathbf{u}$$

$$(13) \quad \mathbf{C} = \mathbf{H}\boldsymbol{\Phi} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} = \lambda_{bt}\mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\zeta}$$

Equation 12 is a SSCRf: it identifies yield responses in discrete candidate management zones to input $x_{k,t}$, while equation 13 is a whole field response function with input by continuous soil test variable interactions. \mathbf{A} includes site-location dummy variables, δ_s . The $\sum_s \delta_s = 0$ restriction on the site-specific indicator variables is useful for testing the SSM hypotheses (Whelan and McBratney, 2000): the t -statistic of the estimated coefficient for a given set of dummy variables indicates whether the intercept, linear, and non-linear terms are significantly different from the average response of the whole field for each of these parameters, and in which direction. The γ_s 's partition the whole field continuous response to inputs and latent soil characteristics (equation 13) into smaller response surfaces corresponding to a given site. In this study, management zones delineated based on phosphorous (P) soil test values. For management purposes, P levels are generally classified into five categories: 0—5 ppm, 6-10 ppm, ..., >20 ppm. Dummy variables indicating these zones are included in the \mathbf{X} matrix. The γ_s coefficient is the marginal effect of the predicted value of the yield for site s conditioned on input $x_{k,t}$, soil test information, and input by soil test information interactions. It is a site-specific weight for the soil test, input, and soil test by input interaction coefficients estimated in (13). This arrangement allows soil test information to enter into EOR input rates an information-based partial budget analysis with respect to site-specific management (SSM).

The matrix $\hat{\mathbf{C}}$ accounts for the latent soil characteristics while avoiding potential multicollinearity problems that may arise in a model that includes data on soil features. $\hat{\mathbf{C}}$ can also be interpreted as a soil fertility index, conditional upon soil test levels, applied inputs, and the interaction between inputs and background soil characteristic levels when included in (12).

System Regression Diagnostics

Baseline regression included estimating the production function using 3SLS and iterated 3SLS (IT3SLS). The null hypothesis that the covariance between equations was zero was rejected using the Breusch-Pagan (1984) Lagrange multiplier test (LM = 2043 and 2042, df = 6, 3SLS and IT3SLS, respectively). To test for the presence of spatial autocorrelation in the least squares residuals, a modified Lagrange multiplier (LM) test for the null hypothesis of no spatial error autocorrelation in the system of equations was conducted (Anselin, 1988, page 147). The system LM (error) statistic is:

$$(14) \quad LM = \mathbf{1}'(\Sigma^{-1} \oplus \mathbf{u}'\mathbf{W}_n\mathbf{u}) \left[tr(\mathbf{W}_n'\mathbf{W}_n)\mathbf{I}_{PT} + tr\mathbf{W}_n'^2\Sigma^{-1} \oplus \Sigma \right]^{-1} (\Sigma^{-1} \oplus \mathbf{u}'\mathbf{W}_n\mathbf{u})' \mathbf{1} \sim \chi^2_{(PT)}$$

where $\mathbf{1}$ is a $PT \times 1$ vector of ones, \oplus is the Hadamard product, and \mathbf{u} is an $n \times PT$ matrix of the least squares residuals. The null hypothesis of no spatial error dependence was rejected at the 1% level for the 3SLS and IT3SLS residuals (LM = 13.14 and 14.93, df = 4, respectively). Based on these results, the equation system was re-estimated using the quasi-ML and GM approaches.

The magnitude of the cross-equation covariance estimates increased when the system was estimated with SAR-QML and GM approaches, however, the magnitude of the cross-equation correlation remained roughly the same between the least squares and SAR estimates (Table 2). The equation-specific squared correlation coefficients were quite high when the system was estimated with the least squares methods (Table 3). Of particular interest for site-specific management is the interaction between inputs and production management zones. The interaction between the dummy variables in equation 12 are significant in four of the five management zones identified in this field, as estimated with IT3SLS (Table 4). The standard errors of the estimates in equation 13 were high, leading to fewer significant parameters (Table 5).

The multicollinearity associated with equation 13 for both periods was mitigated to some extent with the addition of the spatial AR terms, and information about spatial dependence through \mathbf{W} . The frequency of significant parameter estimates for the soil test variables, their interactions with the input variable, and their quadratic terms increased. However, this came at a cost: the predictability of equation 13 appreciably decreased for the soybean equation, dropping from $R^2 = 0.58$ to 0.22 and 0.18 for the SAR-QML and GM approaches, respectively (pseudo- R^2 , as calculated by regressing the predicted values of the dependent variables against the actual values). However, the drop in the R^2 is not surprising: coefficients of determination are usually overstated when errors (time series or otherwise) are positively correlated.

The AR terms in the SAR-QML model were highly significant (Tables 4, 5). The magnitude of the AR terms was greatest for the equation estimating soybean response to soil characteristics and interactions with manure ($\lambda = 0.77$). Estimates for the GM AR terms were similar in magnitude to the SAR-QML λ estimates (Tables 3, 4). Since significance levels for these moment-based AR terms are not identified (Kelejian and Prucha, 1999), standard errors of the GM AR terms are not provided. However, λ are generally considered nuisance parameters, requiring that they only be bounded between -1 and 1.

It is worth noting that the positive sign of the quadratic phosphorous coefficient in the IT3SLS estimate was reversed when the system was re-estimated with SAR-QML or GM. This sign switch has important consequences in the empirical application that follows.

Application of Regression Results

Marginal Analysis and Partial Budget

Fitted crop response functions are often used to make economic comparisons. The following is an example of this application. Whole-field and SSM profitability is determined using a partial budget analysis (Boehlje and Eidman, 1984), and marginal analysis is used to estimate net returns from applied manure (Beattie and Taylor, 1980). That is, profit is maximized when the value of the increased yield from added manure equals the cost of applying an additional gallon; or when the marginal value product equals the marginal factor cost.

For explicit details of the partial budget set-up, corn and soybean prices, input costs, and quasi-fixed costs of information (for example, soil tests), see Lambert, Malzer, and Lowenberg-DeBoer, 2003. Net present value (NPV) of returns to WF and SSM management strategies is estimated as:

$$(15) \quad \max_{(x_{t,k,i,j}, \dots, x_{t,k,i,m}), \dots, (x_{t,k,n,j}, \dots, x_{t,k,n,m})} \sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^M \left(\beta^{t-1} p_t f_t(x_{t,k,i,j}; o_{t,i,j}, w_t) - \beta^t \left(F_t + g_t + v_t + \sum_{k=1}^K r_{t,k} x_{t,k,i,j} \right) \right)$$

where $f(\cdot)$ is a locally concave yield response function; y is output in period t ; x is a vector of managed variable inputs, w is a vector of unmanageable stochastic inputs (e.g. weather); p and r are output and variable input prices; o is a vector of unmanageable, non-stochastic site characteristics (for example, soil depth, soil type, slope and elevation); $i \in N$ and $j \in M$ are site locations; k are variable inputs; g are quasi-fixed costs for information collection (soil test costs); v are quasi-fixed costs for variable rate application (VRA) technology; $\beta^t = (1 + \rho)^{-t}$, ρ is a discount rate ($\rho = 7.5\%$), and p is the price of the crop grown in period t . The discount rate is the estimated return on common stocks over a twenty-year period (1980 – 2000; Chiarella et al., 2002), and F are fixed costs. Solving the first order condition of (15) for x_{NPV}^* gives the NPV-EOR for site s . Substituting x_{NPV}^* into $f(\cdot)$ gives the SSM optimal yield, which is then used to find the net present value of SSM. Equation (15) is adjusted accordingly for WF response analysis.

1. VRM (Variable rate manure): In the first strategy, the producer has site-specific knowledge about corn and soybean yield respond to manure. The producer applies manure at economically optimal rates for each site z . It is assumed that the producer has purchased soil test information to identify management zones, but chooses not to use the information with respect to changing P and K levels, or adjusting pH.
2. VRM-VRF (Variable rate manure – variable rate fertilizer): This scenario considers a producer who simultaneously uses VRA and VRM. Manure and P, K, and lime are applied at EOR levels according to each management zone. The producer is charged a VRA fee (v) for variable application of P, K, and lime, and g for the soil test information.

3. WFM: In this scenario, the producer applies manure uniformly across the WF. An extension agent recommends the rate used. (The recommended rate used here is 3500 gal acre⁻¹).
4. WFM-VRF (Whole-field manure – variable rate fertilizer): The last strategy considers a producer who follows extension recommendations and applies manure at a uniform rate, but manages P, K, and pH using VRA. The producer is charged v and g . Management zones are developed using soil test information.

The second and last scenarios entail maximization of equation 15, changing the levels of P, K, and lime. The remaining scenarios are evaluated at the base soil sample readings for P, K, and lime, and optimal manure rates (as in the case of VRM, solved through the first order conditions of equation 15), or the extension rate of 3500 gal acre⁻¹ (as in the case of WFM).

The NPV for the producer using a VRF strategy (2 and 4) is based on the solution to the optimization problem $\max_{s_{k,z} \in [\underline{s}_{k,z}, \dots, \bar{s}_{k,z}]} NPV$. The soil characteristics in management area ($s_{k,z}$) are bounded between the average of the soil test value for a given characteristic in zone z ($\underline{s}_{k,z}$), and the recommended level corresponding to a target yield level given by Rehm *et al.*, 1994 ($\bar{s}_{k,z}$). This assumes the producer cannot remove nutrient from a given site.

Partial Budget Results

Net present value rankings of the management strategies were identical for the SAR-QML and GM approaches. The rankings of the IT3SLS partial budget results differ from the SAR-QML and GM approaches. Additionally, the IT3SLS NPV estimates for VRM-VRF and WFM-VRF trials were unrealistic, at \$1064 and \$1072 acre⁻¹. These are the scenarios where EOR input levels are updated, conditional upon optimization of equation 15, changing the levels of P, K, and lime. In this scenario based on the IT3SLS estimates, P was binding at the upper level constraint for P. This is expected since the sign of IT3SLS P-quadratic term is positive. On the other hand, if the baseline VRM and WFM strategies can be interpreted as “controls” in this comparison (that is, they are evaluated at the baseline, non-optimized soil test values), then the SAR-QML and GM NPV estimates of VRM-VRF and WFM-VRF are much more believable.

Conclusions

Recently Kelejian and Prucha introduced an approach whereby cross-sectional data modeled as a system of equations could be estimated for spatial error and lag specifications. Earlier, Anselin proposed estimating space-time, or panel data using a SUR approach. This study combines both of these approaches, addressing the space-time nature of panel data, and the endogeneity/simultaneity problems sometimes associated with systems of equations. To estimate the system, Kelejian and Prucha’s GM approach was compared with a QML approach.

To compare the performance of the SAR-QML and GM estimators with the conventional 3SLS estimator, data from an on-farm agronomic trial was used. Based on the results, the 3SLS

estimators perform poorly in the presence of spatial dependence, while the SAR-QML and GM estimators provide good but very similar results. By modeling space-time effects with the SAR models, multicollinearity problems were mitigated, but at the cost of predictability for one of the estimating equations. This loss of predictability is, however, outweighed by the sign-switches observed in some variables, which directly translated into better, more realistic economic analysis. The methodology taken here is easily extended to estimation of panel data sets and simultaneous equations, or other repeated measures data studied in a systems framework. However, a major difference between the GM and QML approaches was computational time: the computational time of the GM approach was about one-eighths that of the QML procedure. For larger systems of equations, the GM approach may be more practical.

Acknowledgments

The authors would like to thank Doug Miller for useful comments and suggestions.

References

- Anselin, Luc. 1988. Spatial Econometrics: Methods and Models. London: Kluwer Academic Publishers.
- Anselin, Luc. 2002. Under the Hood: Issues in the specification and Interpretation of Spatial Regression Models. *Agricultural Economics*, 27:247-267.
- Anselin, Luc, Rodolfo Bongiovanni, and Jess Lowenberg-DeBoer. 2004. A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production. *American Journal of Agricultural Economics*. Forthcoming.
- Aten, Bettina H. 1997. Does space matter? International Comparisons of the Prices of Tradables and Nontradables. *International Regional Science Review* 20(1,2): 35-52.
- Baltagi, Badi H., Seuck Heun Song, and Won Koh. 2003. Testing Panel Data Regression Models with Spatial Error Correlation. *Journal of Econometrics*, 117: 123-150.
- Beattie, B.R. and C.R. Taylor. The Economics of Production. New York: John Wiley & Sons, 1985.
- Bell, Kathleen P., and Nancy E. Bockstael. 2000. Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Micro-level Data. *The Review of Economics and Statistics*, 82(1): 72-82.
- Benirschka, Martin, and James K. Binkley. 1994. Land Price Volatility in a Geographically Dispersed Market. *American Journal of Agricultural Economics*, 76 (May): 185-195.
- Benirschka, Martin, and James K. Binkley. 1995. Optimal Storage and Marketing Over Space and Time. *American Journal of Agricultural Economics*, 77 (August): 512-524.
- Boehlje, M.D., and V.R. Eidman. Farm Management. New York: John Wiley & Sons, 1984.
- Breusch, T.S., and A.R. Pagan. A Lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies* 47, no. 1 (1980): 239-253.
- Bullock, David S., Jess Lowenberg-DeBoer, and Scott Swinton. 2002. Adding value to spatially managed inputs by understanding site-specific yield response. *Agricultural Economics* 1679: 1-13.
- Chiarella, Carl, Willi Semmler, Stefan Mittnik, and Peiyuan Zhu. 2002. "Stock Market, Interest Rate and Output: A Model and Estimation for US Time Series Data" *Studies in Non-linear Dynamics and Econometrics* 6(1): 1-37.

- Chen, Xiaoheng, and Timothy Conley. 2001. A New Semiparametric Spatial Model for Panel Time Series. *Journal of Econometrics*, 105: 59-83.
- Conley, Timothy G. 1999. GMM Estimation with Cross-Sectional Dependence. *Journal of Econometrics*, 92: 1-45.
- Conley, Timothy G., and Giorgio Topa. 2002. Socio-Economic Distance and Spatial Patterns in Unemployment. *Journal of Applied Econometrics*, 17: 303-327.
- Cressie, Noel A.C. 1993. Statistics for Spatial Data. John Wiley & Sons: New York.
- Dillon, John L., and Jock R. Anderson. Analysis of Response in Crop and Livestock Production. New York: Pergamon Press, 1990.
- Druska, Viliam, and William C. Hoorace. 2004. Generalized Moments Estimation for Spatial Panel Data: Indonesian Rice Farming. *American Journal of Agricultural Economics*, 86(1): 185-198.
- Elhorst, J. Paul. 2001. Dynamic Models in Space and Time. *Geographical Analysis*, 33:119-140.
- Florax, Raymond J.G.M., Roelf L. Voortman, and Joost Brouwer. 2002. Spatial Dimensions of Precision Agriculture: a Spatial Econometric Analysis of Millet Yield on Sahelian Coversands. *Agricultural Economics*, 27: 425-443.
- Hardie, Ian W., Tulika A. Narayan, and Bruce L. Gardner. 2001. The Joint Influence of Agricultural and Nonfarm Factors to the Mid-Atlantic Region. *American Journal of Agricultural Economics*, 83(1): 120-132.
- D. F. Heermann, J. Hoeting, S. E. Thompson, H. R. Duke, D. G. Westfall, G. W. Buchleiter, P. Westra, F. B. Peairs, and K. Fleming. 2002. Interdisciplinary Irrigated Precision Farming Research. *Precision Agriculture*, 3, 47-61.
- Gourieroux, C., A. Monfort, and A. Trognon. 1984. Pseudo Maximum Likelihood: Theory. *Econometrica*, 52: 681-700.
- Holloway, Garth, Bhavanni Shankur, and Sanzidur Rahman. 2002. Bayesian Spatial Probit Estimation: a Primer and an Application to HYV Rice Adoption. *Agricultural Economics*, 27: 383-402.
- Kamarianakis, Yiannis. 2003. Spatial-Time Series Modeling: A Review of the Proposed Methodologies. The Regional Economics Application Laboratory (REAL) 03-T-19, May, 2003.
- Giacomini, Rafaella, and Clive W.J. Granger. 2004. Aggregation of Space-Time Processes. *Journal of Econometrics*, 118: 1-26.

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

Klejjan, Harry H., and Ingmar R. Prucha. 2004. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *Journal of Econometrics*, 118: 27-50.

Lambert, Dayton, Gary L. Malzer, and Jess Lowenberg-DeBoer. 2003. A Systems Approach Incorporating Soil Test Information into Site-Specific Manure Management Recommendations. Staff Paper #03-13, Purdue University Department of Agricultural Economics, West Lafayette, IN, 47906, <http://agecon.lib.umn.edu/cgi-bin/view.pl>.

Lence, Sergio H., and Ashok K. Mishra. 2003. The Impacts of Different Farm Programs on Cash Rents. *American Journal of Agricultural Economics*, 85(3): 753-761.

Leonard, Michael, and Renne Somy. 1997. *Forecasting Geographic Data*. The SAS Institute, Inc., Cary, North Carolina.

LeSage, J.P., and Zheng Pan. 1995. Using Spatial Contiguity as Bayesian Prior Information in Regional Forecasting Models. *International Regional Science Review*, 18(1):33-53.

Mittelhammer, Ron C., George G. Judge, and Douglas J. Miller. 2000. Econometric Foundations. Cambridge University Press: Cambridge.

Ord, Keith. 1975. Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Society*. 70:120-126.

Munroe, Darla, Jane Southworth, and Catherine M. Tucker. 2002. The Dynamics of Land-cover Change in Western Honduras: Exploring Spatial and Temporal Complexity. *Agricultural Economics* 27:355–369.

Pace, R. Kelley, Ronald Berry, Otis W. Gilley, and C.F. Sirmans. 2000. A Method for Spatial-Temporal Forecasting with and Application to Real Estate Prices. *International Journal of Forecasting*, 16: 229-246.

Pinske, Joris, and Margaret E. Slade. 1998. Contracting in Space: An Application of Spatial Statistics to Discrete-choice Models. *Journal of Econometrics*, 85: 125-154.

Rehm, George, Michael Schmitt, and Robert Munter. 1994. Fertilizer Recommendations for Agronomic Crops in Minnesota. *Minnesota Extension Bulletin BU-6240-E*.

Roe, Brian, Elena G. Irwin, and Jeff S. Sharp. 2002. Pigs in Space: Modeling the Spatial Structure of Hog Production in Traditional and Nontraditional Production Regions. *American Journal of Agricultural Economics*, 84(2): 259-278.

Silveira-Neto, Raul M., and Carlos R. Azzoni. 2003. Location Spillovers and Growth Among Brazilian States. TD Nereus (The University of São Paulo Regional and Urban Economics Lab) 11-2003.

Swinton, Scott. 2002. Capturing Household-level Spatial Influence in Agricultural Management using Random Effects Regression. *Agricultural Economics*, 27: 371-381.

Whelan, B.M., and A.B. McBratney. 2000. The “null hypothesis” of precision agriculture management. *Precision Agriculture*, 2: 265-279.

Table 1. Notation used to identify corn and soybean response functions (equations 12, 13).

\mathbf{y} is a geo-referenced $n \times 1$ vector of yield unit ⁻¹ for the crops in each study;
\mathbf{X} is an $n \times k$ matrix of fixed effects: input levels unit ⁻¹ , their squares, and location dummy variables for field sites s , $s \in [1, \dots, S]$ sites;
$\hat{\mathbf{C}}$ is an $n \times s$ vector of predicted yield values;
\mathbf{H} is an $n \times (q + k - s)$ matrix of instruments including the q -th soil test characteristic and k input by soil characteristic interactions;
\mathbf{W} is an $n \times n$ matrix of spatial or temporal relations between observations following the queen contiguity criterion;
\mathbf{A} is a $k \times 1$ vector of fixed effects parameters to be estimated, including location dummy variables (δ_s) to identify site-specific response constrained as $\sum_s \delta_s = 0$;
Φ is a $(q + k - s) \times 1$ matrix of linear and quadratic coefficients for a continuous yield response function;
j -th soil characteristic managed as an input (for example, K and P);
γ is a $s \times 1$ matrix of site locations;
λ_{at} and λ_{bt} are spatial AR parameters for the a -th and b -th equation, $a = 1, 2$, $b = 1, 2$, $t = 1, 2$;
ϵ and ζ are $n \times 1$ vectors of (possibly) autocorrelated disturbances;
\mathbf{u} and \mathbf{e} are $n \times 1$ vectors of i.i.d. disturbances $\sim N(\mathbf{0}, \sigma^2 \mathbf{I})$.

Table 2. Covariance and correlation matrices estimated with least squares, SAR-QML, and SAR-GM.

	Covariance Matrix, $\Sigma(\Lambda)$				Cross-equation correlation		
	Corn, Eq. 1	Corn, Eq. 2	Soybean, Eq. 1	Soybean, Eq. 2	Corn, Eq. 2	Soybean, Eq. 1	Soybean, Eq. 2
3SLS							
Corn, Eq. 1	68.57	-67.03	9.73	-9.63	-0.52	0.91	-0.27
Corn, Eq. 2	-67.03	240.27	-9.42	28.22	1.00	-0.47	0.42
Soybean, Eq. 1	9.73	-9.42	1.68	-2.11		1.00	-0.37
Soybean, Eq. 2	-9.63	28.22	-2.11	19.16			1.00
IT3SLS							
Corn, Eq. 1	99.89	-68.81	16.60	-10.17	-0.44	0.93	-0.23
Corn, Eq. 2	-68.81	241.30	-13.76	28.56	1.00	-0.50	0.42
Soybean, Eq. 1	16.60	-13.76	3.17	-3.14		1.00	-0.40
Soybean, Eq. 2	-10.17	28.56	-3.14	19.38			1.00
SAR-QML							
Corn, Eq. 1	102.27	-71.62	16.54	-10.44	-0.45	0.93	-0.22
Corn, Eq. 2	-71.62	244.30	-13.83	32.10	1.00	-0.50	0.44
Soybean, Eq. 1	16.54	-13.83	3.08	-3.13		1.00	-0.39
Soybean, Eq. 2	-10.44	32.10	-3.13	21.41			1.00
SAR-GM							
Corn, Eq. 1	104.88	-75.25	17.03	-10.70	-0.47	0.92	-0.22
Corn, Eq. 2	-75.25	244.27	-14.36	32.84	1.00	-0.51	0.44
Soybean, Eq. 1	17.03	-14.36	3.24	-3.45		1.00	-0.41
Soybean, Eq. 2	-10.70	32.84	-3.45	22.34			1.00

Table 3. Model fit statistics for each equation estimated with least squares, SAR-QML, and SAR-Gm approaches.

	Corn, Eq. 1	Corn, Eq. 2	Soybean, Eq. 1	Soybean, Eq. 2
3SLS	0.95*	0.84	0.95	0.58
IT3SLS	0.95	0.84	0.95	0.58
SAR-QML	0.95	0.84	0.95	0.22
SAR-GM	0.94	0.84	0.95	0.18

*Squared correlation coefficient.

Table 4. Iterated least squares, SAR-QML, and SAR-GM estimates, equation 12.

Variable	IT3SLS				SAR-QML				SAR-GM			
	<u>Corn</u>		<u>Soybean</u>		<u>Corn</u>		<u>Soybean</u>		<u>Corn</u>		<u>Soybean</u>	
	Estimate	T	Estimate	T	Estimate	Z	Estimate	Z	Estimate	Z	Estimate	Z
Intercept	21.32	5.86	0.01	0.01	18.76	14.33	-0.72	-1.98	17.83	6.63	-1.36	-2.22
G*	-0.01	-4.57	-1.26E-03	-5.81	-0.01	-10.39	-1.26E-03	-12.11	-0.01	-6.65	-1.32E-03	-9.84
G ²	7.33E-07	4.80	1.24E-07	4.95	7.68E-07	10.67	1.25E-07	10.35	7.96E-07	6.85	1.28E-07	8.36
d ₂ **	-32.18	-11.09	-4.28	-8.35	-32.02	-26.23	-4.29	-20.25	-31.96	-15.32	-4.31	-15.04
d ₃	-2.30	-0.74	-0.04	-0.07	-1.29	-0.90	-0.09	-0.38	-1.57	-0.65	-0.10	-0.33
d ₄	13.40	2.77	0.47	0.58	14.55	6.87	0.66	1.85	14.06	4.05	0.61	1.33
d ₅	4.33	0.64	-3.66	-3.05	2.16	0.72	-3.15	-6.08	2.70	0.50	-3.12	-4.34
G x d ₂	0.02	14.00	2.56E-03	12.02	0.02	28.17	2.54E-03	23.60	0.02	19.34	2.53E-03	19.23
G x d ₃	-7.00E-05	-0.05	-3.20E-04	-1.39	1.79E-04	0.24	-2.80E-04	-2.30	5.17E-05	0.05	-2.99E-04	-2.01
G x d ₄	-0.02	-6.67	-3.13E-03	-7.56	-0.02	-14.67	-3.03E-03	-16.11	-0.02	-9.78	-3.03E-03	-12.79
G x d ₅	-0.01	-4.49	-1.56E-03	-4.56	-0.01	-9.17	-1.45E-03	-8.92	-0.01	-6.09	-1.45E-03	-7.15
G ² x d ₂	-1.96E-06	-12.95	-2.82E-07	-11.13	-1.99E-06	-26.66	-2.78E-07	-22.22	-1.98E-06	-17.88	-2.77E-07	-18.00
G ² x d ₃	-7.99E-08	-0.48	5.84E-08	2.11	-1.13E-07	-1.32	5.10E-08	3.58	-1.06E-07	-0.86	5.31E-08	3.04
G ² x d ₄	2.07E-06	7.00	4.31E-07	8.71	2.06E-06	15.13	4.13E-07	18.17	2.05E-06	9.91	4.14E-07	14.41
G ² x d ₅	1.49E-06	6.13	2.33E-07	5.74	1.44E-06	12.38	2.19E-07	11.23	1.45E-06	8.36	2.21E-07	9.08
Γ x d ₁	0.85	17.05	0.91	28.00	0.86	52.15	0.94	84.36	0.88	25.62	0.95	54.47
Γ x d ₂	0.96	32.36	1.06	43.52	0.98	88.81	1.08	112.11	0.99	41.56	1.09	70.17
Γ x d ₃	0.95	35.02	1.04	47.78	0.96	98.23	1.06	118.90	0.98	42.47	1.08	74.39
Γ x d ₄	0.94	24.17	1.07	39.14	0.94	65.30	1.08	101.06	0.96	31.13	1.09	63.58
Γ x d ₅	0.91	18.46	1.12	35.11	0.94	48.52	1.13	96.00	0.94	24.80	1.14	61.03
λ					0.10	236.87	0.10	243.38	0.04		0.06	

*Gallons acre⁻¹ liquid hog manure; **Management zone indicator.

Table 5. Iterated least squares, SAR-QML, and SAR-GM estimates, equation 13.

Variable	IT3SLS				SAR-QML				SAR-GM			
	Corn		Soybean		Corn		Soybean		Corn		Soybean	
	Estimate	T	Estimate	T	Estimate	Z	Estimate	Z	Estimate	Z	Estimate	Z
Intercept	-245.69	-0.37	86.22	0.47	-485.47	-3.03	-146.23	-15.89	-487.03	-3.85	-178.12	-19.81
G	0.02	3.16	1.72E-03	0.89	0.02	10.06	-6.39E-04	-4.55	0.02	13.04	-1.11E-03	-8.12
P	29.70	2.04	2.21	0.55	35.43	9.08	8.04	32.37	36.00	11.45	8.82	36.11
Om	-86.34	-0.64	-12.74	-0.34	-92.03	-2.59	16.17	7.75	-87.34	-3.10	20.43	10.02
Zn	-138.14	-0.69	-42.69	-0.77	-113.81	-2.47	-31.58	-11.41	-115.48	-3.15	-28.59	-10.57
K	1.51	0.84	0.26	0.53	1.92	3.96	1.09	35.15	1.96	4.99	1.28	42.39
pH	129.21	0.47	-12.29	-0.16	190.64	2.80	25.93	6.47	185.70	3.43	28.08	7.18
G ²	-2.76E-06	-20.64	-4.47E-07	-11.77	-2.58E-06	-75.68	-3.02E-07	-139.13	-2.55E-06	-94.71	-2.82E-07	-131.72
G x P	-3.30E-04	-3.09	-4.00E-05	-1.33	-3.84E-04	-11.37	-9.19E-05	-42.20	-3.87E-04	-14.19	-8.90E-05	-41.69
G x OM	-1.19E-03	-1.73	1.45E-04	0.76	-7.76E-04	-3.12	3.95E-04	25.32	-8.28E-04	-4.11	3.84E-04	25.11
G x Zn	-1.11E-03	-0.85	-8.10E-04	-2.20	-1.01E-03	-2.64	-2.96E-04	-12.76	-9.43E-04	-3.09	-2.41E-04	-10.66
G x K	-2.00E-05	-1.32	-1.24E-06	-0.29	2.15E-06	0.42	2.17E-05	65.46	7.19E-07	0.17	2.27E-05	70.15
G x pH	4.04E-03	2.93	7.78E-04	2.03	2.77E-03	6.17	4.61E-05	1.67	2.68E-03	7.42	5.78E-05	2.15
P ²	-3.62E-03	-0.03	0.07	1.85	-0.08	-2.46	-0.03	-14.04	-0.08	-3.33	-0.04	-20.17
OM x P	-5.71	-2.43	-1.41	-2.13	-4.92	-8.53	-0.67	-18.95	-4.89	-10.68	-0.76	-22.14
Zn x P	7.14	2.25	0.45	0.51	6.51	9.00	0.36	7.86	6.61	11.38	0.63	14.18
P x K	0.01	0.50	-0.01	-1.09	0.03	3.90	0.01	29.36	0.03	4.88	0.01	37.02
pH x P	-2.98	-1.27	0.57	0.86	-4.48	-6.93	-1.15	-28.57	-4.61	-8.87	-1.33	-33.64
OM ²	8.11	1.33	3.17	1.85	6.16	3.52	0.37	3.85	6.10	4.45	0.26	2.75
Zn x OM	24.85	0.91	5.05	0.66	26.26	3.34	5.49	12.36	25.43	4.11	5.96	13.73
OM x K	0.56	2.51	0.12	1.94	0.51	8.26	0.13	32.80	0.51	10.20	0.14	36.29
OM x pH	-6.35	-0.26	-4.66	-0.69	-3.29	-0.53	-7.91	-22.44	-3.77	-0.77	-8.73	-25.40
Zn ²	-34.96	-1.57	-0.59	-0.10	-34.21	-5.73	-2.54	-7.28	-33.89	-7.16	-3.89	-11.41
Zn x K	-1.23	-2.59	-0.14	-1.02	-1.38	-11.81	-0.47	-67.54	-1.40	-15.01	-0.54	-78.86
Zn x pH	48.60	1.25	7.91	0.73	48.98	5.64	16.80	31.78	49.76	7.18	18.06	34.96
K ²	1.26E-03	0.46	-3.00E-05	-0.04	1.98E-03	2.61	1.14E-03	26.09	1.93E-03	3.22	1.04E-03	24.64
pH x K	-0.44	-1.40	-0.08	-0.95	-0.51	-6.01	-0.27	-53.22	-0.51	-7.51	-0.30	-59.63
pH ²	-8.43	-0.29	2.62	0.32	-12.26	-1.65	3.00	6.84	-11.66	-1.98	3.34	7.81
λ					0.16	205.71	0.75	248.28	0.19		0.77	

Table 6. Comparison of NPV estimated with IT3SLS, SAR-QML, and SAR-GM.

	Rank	Estimate -----\$ acre ⁻¹ -----
IT3SLS		
VRM [‡]	3	576.68
VRM-VRF**	2	1064.26
WFM [†]	4	569.39
WFM-VRF	1	1071.67
SAR-QML		
VRM	3	582.50
VRM-VRF	1	600.95
WFM	4	574.74
WFM-VRF	2	597.36
SAR-GM		
VRM	3	582.80
VRM-VRF	1	597.47
WFM	4	575.21
WFM-VRF	2	592.56

[†]WF is evaluated at the uniform (recommended) rate of 3500 gal acre⁻¹.

[‡] Variable rate manure is evaluated at the baseline soil test readings, and the economically optimal input rate from the solution of equation 15.

**VRF scenarios entail optimization of equation 15, with the choice variables as the soil characteristics, P, K, and lime (as pH).