The Spatial Structure of Farmland Values: A Semiparametric Approach

Haoying Wang

Email: halking@psu.edu
Web: https://sites.google.com/site/halkingwang/

Department of Agricultural Economics, Sociology and Education
The Pennsylvania State University, University Park, PA, 16802-5600

Current Draft: June 3, 2013

Selected Poster prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting
Washington, DC, August 4-6, 2013.

The Spatial Structure of Farmland Values: A Semiparametric Approach

Haoying Wang
Email: halking@psu.edu
Web: https://sites.google.com/site/halkingwang/
Department of Agricultural Economics, Sociology and Education
The Pennsylvania State University, University Park, PA, 16802-5600

Abstract

Although accounting for the spatial - temporal relationships in farmland valuation has gained attention in the literature recently, misspecification and incorrectly imposed assumptions on spatial weighting matrix can often produce misleading estimates and inference compared to maintaining ignorance of spatial dependence structure among spatially observed farmland values. In this study I assemble a panel data set using Pennsylvania county level farmland values reported in the U.S. Census of Agriculture between 1982 and 2007, and estimate the spatial weighting matrix among farmland values semiparametrically. A spatial lag panel data model with the consistently estimated spatial weighting matrix is then estimated via maximum likelihood estimation (MLE). The results show that the proposed approach can substantially improve the goodness of fit of the spatial hedonic model of farmland values therefore the reliability of obtained price elasticity estimates.

Introduction

• Farmland Value - commonly refer to the value of a farm real estate (land and structures on it) package.
• Why important? Farmland accounted for **85%** of the total value of U.S. farm assets in 2010 (ERS, USDA); **3/4** of the total assets value comes from farm real estate, while it accounts for only about **half** of the total-sector debt (Moss and Schmitz, 2008); nationwide farmland value increase over last three decades;
• Why care? Important property tax base; important input in all sorts of integrated assessments (e.g., climate change, ecosystem conservation, urban system planning).

• **Present Value Models** v.s. **Hedonic Pricing Models**: different suggestions for the potential explanatory variables to include, but similar reduced form in empirical estimation;
• Determinants Identified: productivity and profitability, average farm size, inflation, environmental vulnerability, land use policy, **spatial location**, land price volatility, local demographics;
• Non-Spatial Method: (Sandrey et al., 1982; Mendelsohn et al., 1994; Boisvert et al., 1997; Moss, 1997);
• Spatial Method: (Huang et al., 2006; Schlenker et al., 2006; Jeanty et al., 2010; Brady and Irwin, 2011) (1) - pre-specified spatial weighting scheme; (2) - incorrect post-regression (output) analysis.

Econometric Model

Hedonic model of farmland values: $p_{it} = \alpha_i + X_{it}\beta + \varepsilon_{it}$, $\varepsilon_{it} \sim N(0, \sigma^2)$
Spatial hedonic model with **endogenous interactions**:

$$p = \rho(I_T \otimes W_N)p + X\beta + \varepsilon$$

with panel data, it can be written as a **fixed effects** spatial lag model:

$$p = \rho(I_T \otimes W_N)p + (\iota_T \otimes \alpha) + X\beta + \varepsilon$$

where T is the # of time periods, N is the # of cross-section units, ι_T is a vector of T ones. p is farmland value, X is the matrix of explanatory variables, and W_N is a spatial weighting matrix.

In the literature, the spatial weighting matrix tends to be pre-specified based on some discrete or continuous distance measure (e.g., the binary contiguity matrix). In this paper, I use Pinkse et al. (2002)'s semiparametric approach to estimate the spatial weighting matrix first. By substituting the consistently estimated weighting matrix into the spatial panel data models shown above, I then estimate the coefficients (elasticity) for the determinants of farmland values via maximum likelihood estimation (MLE).

Data

The U.S. Census of Agriculture, 1982 - 2007;
Climatic data from NOAA, monthly weather station data;
Housing Price Index (HPI) data from FHFA, MSA level monthly data;
PA Agricultural Statistics Annual Statistical Bulletin, 1982 - 2007.

Estimation: Stage 1

• Following Pinkse et al. (2002), in a cross-sectional setting, the spatial hedonic model can be written as: $p_i = \sum_{j \neq i} g(d_{ij})\, p_j + x_i\beta + \varepsilon_i$, where $g(\cdot)$ is a function of some distance measures, d_{ij}.
• Using semiparametric series expansion: $g(d) = \sum_{l=1}^{\infty} \gamma_l e_l(d)$, where γ_l's are unknown coefficients to be estimated, and e_l's form a basis of the function space to which $g(\cdot)$ belongs, then we have:

$$p_i = \sum_{l=1}^{k_n} \gamma_l \sum_{j \neq i} e_l(d_{ij}) p_j + \sum_{l=k_n+1}^{\infty} \gamma_l \sum_{j \neq i} e_l(d_{ij}) p_j + x_i\beta + \varepsilon_i$$

• Define new error term as: $\xi_i = \sum_{l=k_n+1}^{\infty} \gamma_l \sum_{j \neq i} e_l(d_{ij}) p_j + \varepsilon_i$, and a matrix Z with (i, l) element as $\sum_{j \neq i} e_l(d_{ij}) p_j$, then we have:

$$p = Z\gamma + X\beta + \xi$$

where k_n is the number of terms in the finite part of expansion.

• Note that price variable p_j on the right hand side are endogenous, let q_j be the instrument for p_j, and Q be the matrix of instruments formed by $\sum_{j \neq i} e_l(d_{ij}) q_j$ and be independent of ξ_i, and $P_Q = Q(Q^T Q)^{-1} Q^T$ be the orthogonal projection matrix onto the columns of Q, we have:

$$P_Q p = P_Q Z\gamma + P_Q X\beta + P_Q \xi$$

• Further, we assume instrument matrix Q is independent of X. Here we are only interested in estimating γ_l's, the above estimation can be then simplified as (ξ' is the new error term):

$$P_Q p = P_Q Z\gamma + + P_Q \xi'$$

• Now the estimates of γ_l's can be obtained by classic IV estimator:

$$\hat{\gamma} = (Z^T P_Q Z)^{-1} Z^T P_Q p$$

then $g(\cdot)$ is estimated as: $\hat{g}(d) = \sum_{l=1}^{k_n} \hat{\gamma}_l e_l(d)$.

• In stage 1, the **endogeneity** issue is addressed via IV estimation. An advantage of the **independence** assumption between Q and X and above simplification is that, the estimates of γ and estimated spatial weighting matrix are **exogenous** to the spatial hedonic model in the second stage where X is the matrix of exogenous regressors.
• $\hat{g}(d_{ij})$ thus gives an estimate of the (i, j) element of spatial weighting matrix W_N, which represents an empirical measure on the interaction of farmland values between county i and j given $i \neq j$.

Results: Model Comparison

Model	GDP deflated Farmland Values			PPI deflated Farmland Values		
W_N	$W^{(0,1)}$	W^A	$\widehat{W}$	$W^{(0,1)}$	W^A	$\widehat{W}$
# of obs	384	384	384	384	384	384
ρ (s. e.)	0.5819 (0.0421)	0.5860 (0.0403)	0.0500 (0.0155)	0.7250 (0.0317)	0.7130 (0.0308)	0.6820 (0.0303)
Log Likelihood	-332.0093	-327.8612	-258.6543	-309.9263	-300.3449	-289.3106

$W^{(0,1)}$: binary contiguity matrix; $\widehat{W}$: estimated weighting matrix;
W^A: binary contiguity matrix weighted by normalized farmland area.

Estimation: Stage 2

• In Stage 1, the distance measure d_{ij} is defined in following way:

$$d_{ij} = A_j \mathrm{W}_{ij} \;\; and \;\; \mathrm{W}_{ij} = \begin{cases} 1, & if\ county\ i\ and\ j\ share\ border \\ 0, & otherwise \end{cases}$$

where A_j is the normalized (to [0,1]) total farmland area (average over 6 census years) in county j.
• Substituting in the estimated spatial weighting matrix from Stage 1 after row standardization, the spatial hedonic model becomes:

$$p = \rho\left(I_T \otimes \widehat{W_N}\right)p + (\iota_T \otimes \alpha) + X\beta + \varepsilon$$

which is estimated via MLE with the concentrated log-likelihood function proposed by Elhorst (2010). The inconsistency issue of σ^2 estimate is corrected by the data transformation proposed by Lee and Yu (2010).
• Assuming normally distributed error term ε, the concentrated log-likelihood function is given by:

$$LogL = -\frac{NT}{2}\log(2\pi\sigma^2) + T log\left|I_N - \rho\widehat{W_N}\right| - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(p^*_{it} - \rho\sum_{j=1}^{N}\widehat{W}_{ij}p^*_{it} - X^*_{it}\beta\right)^2$$

where p^*_{it} and X^*_{it} are given by following transformation: $p^*_{it} = p_{it} - \frac{1}{T}\sum_{t=1}^{T} p_{it}$, $X^*_{it} = X_{it} - \frac{1}{T}\sum_{t=1}^{T} X_{it}$

Results: Estimates

Model	ln(gdp_price)			
Marginal Effect	$\hat{\beta}$	Total	Direct	Indirect
farm area(%)	-0.2142	-0.2250	-0.2136	-0.0114
# of farms	0.3699	0.3887	0.3691	0.0196
farm size	-0.1575	-0.1661	-0.1577	-0.0084
asset return($)	0.1027**	0.1080	0.1026	0.0054
full owner(%)	0.2021***	0.2125	0.2018	0.0107
tenant owner(%)	-0.0115	-0.0120	-0.0114	-0.0006
Investment($)	0.2802***	0.2947	0.2798	0.0149
hired labor($)	-0.0560**	-0.0589	-0.0559	-0.0030
operator age	3.2874***	3.4578	3.2833	0.1745
dairy farm(%)	0.0191	0.0201	0.0191	0.0010
dpnp(0.1cm)	-0.0112***	-0.0117	-0.0111	-0.0006
dpnt(0.1F)	0.0194***	0.0204	0.0194	0.0010
mntm(F)	-0.2117	-0.2226	-0.2114	-0.0112
tpcp(inch)	0.0867	0.0913	0.0867	0.0046
q4	0.0217**	0.0228	0.0216	0.0012
q8	-0.0209***	-0.0220	-0.0209	-0.0011

Climatic variables: dpnp, dpnt, mntm, tpcp; Price uncertainty variables: q4, q8;
All explanatory variables are logarithm transformed.

Conclusion

• **With two different approaches to deflating farmland values, our models with the estimated spatial weighting matrix performs better than models with the pre-specified weighting matrices in terms of goodness of fit.**

• **Return to assets, farm ownership, distance to metro area, price uncertainty, investment level, and productivity are among the key determinants of farmland values.**

• **I only see weak evidence of climate change impacts on farmland values. The results indicate that short term climate variability and long term climate change have different impacts on farmland values.**

• **There are strong interactions among farmland markets across regions, which well explains the simultaneous pattern of farmland values change we have observed across regions.**

Selected References

• **Elhorst, J. P.** 2010. Spatial Panel Data Models. In Fischer, M. M., Getis, A. (Eds.) Handbook of Applied Spatial Analysis: 377-407. Springer: Berlin Heidelberg, New York, U.S.A.
• **Huang, H., G. Y. Miller, B. J. Sherrick, and M. I. Gomez.** 2006. "Factors Influencing Illinois Farmland Values." American Journal of Agricultural Economics 88(2): 458-470.
• **Lee, L. F., and J. Yu.** 2010a. "Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects." Journal of Econometrics 154(2): 165-185.
• **Moss, C. B.** 1997. "Returns, Interest Rates, and Inflation: How They Explain Changes in Farmland Values." American Journal of Agricultural Economics 79(4): 1311-1318.
• **Pinkse, J., and M. E. Slade.** 2010. "The Future Of Spatial Econometrics." Journal of Regional Science 50(1): 103-117.
• **Pinkse, J., M. E. Slade., and C. Brett.** 2002. "Spatial Price Competition: A Semiparametric Approach." Econometrica 70(3): 1111-1153.