Journal of Agricultural and Resource Economics 49(1):63–80 ISSN: 1068-5502 (Print); 2327-8285 (Online)

doi: 10.22004/ag.econ.337550

Spatial and Temporal Spillovers in US Cropland Values

J. Wesley Burnett, Donald J. Lacombe, and Steven Wallander

While markets for agricultural commodities are often global, markets for cropland can be highly localized. Soil makeup and climate also tend to make farmland rental markets localized. This study compares several econometric models to measure the extent of spatial spillovers within these markets. A fully dynamic spatial model estimated on cropland transactions shows that sales are highly localized in the short term but face greater neighboring spillovers in the long term. Accounting for longer-term spatial interdependence of farmland markets can help to improve land value forecasts.

Key words: agricultural land values, capitalization model, farmland cash rental rates, spatiotemporal econometrics

Introduction

Land is one of the primary assets in crop production, making up 83% of total farm assets (Burns et al., 2018). Moreover, farmland is one of the primary assets for farmer wealth and the leading source of collateral for agricultural loans (Nickerson et al., 2012). Agricultural land is either owned or rented under a variety of lease contract types and levels of complexity. The most prominent types of lease agreements include cash, share, and hybrid contracts (Bigelow, Borchers, and Hubbs, 2016). The cost of access to farmland offers policy implications for domestic agriculture, and previous research has shown that land costs can impact operational decisions, including the adoption of on-farm conservation practices (Soule, Tegene, and Wiebe, 2000). Today, approximately 40%, or 355 million acres, of farmland in the lower 48 United States is rented (Bigelow, Borchers, and Hubbs, 2016).

One of the difficulties of analyzing farmland values is that markets for farmland are “thin." In other words, there are relatively few buyers and sellers, and—consequently—sales transactions take place infrequently (relative to residential transactions). According to Bigelow and Hubbs (2016), less than 4% of all US land in farms was expected to be sold from 2014 to 2019, and just over 2% of farmland was expected to be sold to a nonrelative of the current owner. According to Johnson (2010), the general rule of thumb is that approximately 3% of agricultural land turns over each year.

Another challenge in analyzing farmland values is that, generally, only proxy variables are available to the researcher to separate agricultural land values and nonagricultural values (Robison and Koenig, 1992). Farm returns and interest rates are the main determinants of land values (Burns et al., 2018); however, data on operator-level farmland returns are not readily available. The USDA

J. Wesley Burnett (corresponding author; wesley.burnett@usda.gov) is a research economist and Steven Wallander (steven.wallander@usda.gov)is an economist in the Economic Research Service at the US Department of Agriculture. Donald J. Lacombe (donald.lacombe@ttu.edu) is an associate professor in the School of Financial Planning at Texas Tech University.
The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or US Government determination or policy. This research was supported by the intramural research program of the USDA’s Economic Research Service and Texas Tech University. Donald Lacombe received extramural funding from the Economic Research Service in support of this research project.

Review coordinated by Jeffrey J. Reimer.

Census of Agriculture offers numerous proxies for farm returns, but the data are only available every 5 years. The US Bureau of Economic Analysis (2019) offers annual (intercensal) estimates of agricultural cash receipts from marketing (a proxy for returns) based on agricultural census data and other macroeconomic accounts.[1] Although the Bureau's measure of returns are offered annually, the data are aggregated; operator-level cash receipts are not necessarily directly observed.

Another potential proxy for operator returns is cash rental rates. The rental rate proxy is based on the findings of Kirwan (2009), who argued that agricultural rental rates respond primarily to innovations in expected farm returns in the current period.[2] Moreover, Gutierrez, Westerlun, and Erickson (2007) found a relationship between farmland values and cash rents; their estimates suggested that a 1% increase in cash rents is associated with a 1% increase in farmland prices. The county-level farmland rental rate data used in the current study are collected by the National Agricultural Statistics Service annually. Thus, the rental rate data offer relatively fine variation across time and space.

Against this backdrop, we analyze farmland market values as a function of farm characteristics, geographic and year fixed effects, and neighboring land prices. This paper offers at least three contributions to the literature: First, The study uses private, agricultural sales transaction data as opposed to survey elicited prices (e.g., USDA survey on farmland values).[3] Other studies have examined the relationship across data sources (e.g., the USDA's June Area Survey and market transactions, see Kuethe and Ifft, 2013; Bigelow, Ifft, and Kuethe, 2020). For example, Bigelow, Ifft, and Kuethe (2020) found that the separate data sources provided similar weight-value distributions, but regression estimates identified differences in marginal effects estimates that could be attributed to market thinness. Despite this potential shortcoming, we use the transaction data because they allow us to examine county-level average farmland values, while the June Area Survey only reports data aggregated to the state level.

Second, we use a spatiotemporal econometric approach to capture the underlying spatial and temporal dynamics within agricultural land markets. The estimated spatial autocorrelation parameters are found to be positive and highly statistically significant across all model specifications, implying spatial dependence among local and neighboring land values. The findings suggest that nonirrigated cash rental rates—as a proxy for annual agricultural returns—are positively correlated with farmland values in the short and the long run. Additionally, estimation results suggest that, on average, county-level Conservation Reserve Program (CRP) participation is negatively correlated (in the short run) with agricultural land values over the 2008–2020 sample period.

Three, we provide out-of-sample forecasts for each of the models and find that a dynamic spatial panel data model (the preferred model) provides the lowest forecasting errors.

Background

Figure 1 depicts the national trend in average US farm real estates values between 1970 and 2020 (Callahan, 2020). The trend in nominal values suggests that per acre farmland values grew on average by about 6%/year over this period. At this rate of growth, the price of farmland would be expected to double nearly every 12 years.

[1] According to the Bureau of Economic Analysis, cash receipts from marketing consist of gross revenue received by farmers from the sale of crops, livestock, and the value of defaulted loans made by the Commodity Credit Corporation.

[2] Kirwan's (2009) findings are echoed by survey analyses conducted by the Agricultural Extension Office of the University of Nebraska-Lincoln, which found that per acre rental rates are determined, in part, by expected agricultural returns (Vyhnalek et al., 2020).

[3] Zakrewicz, Brorsen, and Briggeman (2012) argued that USDA land values are reported as representing the prices in January; as such, the prices have a tendency to more closely represent the land values in the first and second quarters of each year.

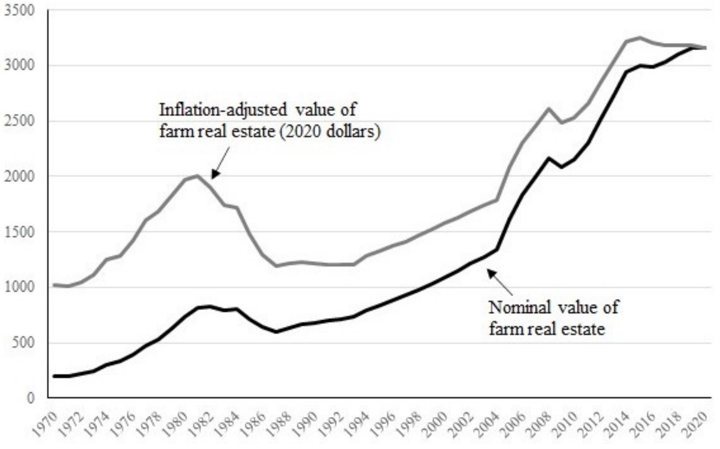

Figure 1. Average Values of US Farm Real Estate ($ /acre), 1970–2020

Notes: The grey line illustrates the inflation-adjusted value of farm real estate values in 2020 US dollars per acre. The black line depicts the nominal value of farm real estate values per acre.
Source: Callahan (2020).

Literature Review

There is a rich past literature of analysis of agricultural land values. The most common approach to modeling farmland prices is based on the land capitalization model approach (Melichar, 1979; Alston, 1986; Burt, 1986; Brueckner, 1990; Plantinga, Lubowski, and Stavins, 2002; Nickerson and Zhang, 2014; Kuethe and Oppedahl, 2021). Generally speaking, the capitalization approach (discussed in further detail in the next section) models the current value of farmland as a function of future agricultural net returns conditional on other codeterminants. It is a flexible approach that incorporates returns and the opportunity cost of capital. Examples include Alston (1986), who found that the change in real land price growth is explained by real growth in net returns. Burt (1986) found support for the capitalization model using a relatively sophisticated econometric assessment (autoregressive moving average). More recently, Kuethe and Oppedahl (2021) used a similar capitalization model, but their empirical analysis examined farmland prices based on a survey of agricultural bankers.

Over the past 2 decades, increasingly sophisticated regressions and refinements to the original model have been developed to account for spatial dependence (spatial autocorrelation) within the underlying farmland price data. Using a spatial econometric model approach, Benirschka and Binkley (1994) found increasing land price variation with distance to market. Moreover, Huang et al. (2006) used a spatial econometric model to examine farmland values in Illinois. They found that farmland values decreased with distance to Chicago and other large cities and increased with population density and personal income. Dillard et al. (2013) used a spatial econometric model to examine farm real estate prices in Indiana from 2003 to 2006. Similar to the current study, the authors estimated spatial autocorrelation coefficients and then decomposed the model's estimates into direct, indirect, and total marginal effects.

Beyond empirical studies of land capitalization, a new emerging literature focuses on nationwide land valuation analyses. In part, the motivation of this literature is to create a "land" natural capital account to supplement existing capital accounts within a country's estimate of gross domestic product (Wentland et al., 2020). A recent contribution to this literature includes high-resolution mapping of estimated values of private lands in order to predict the cost of land conservation interventions in the contiguous United States (Nolte, 2020). The current study contributes to this literature by offering a methodology for estimating the value of agricultural lands. As noted above, past attempts to evaluate farmlands have been challenging because private transactions of agricultural lands are so thin or infrequent (relative to residential transactions).

Domestic Agricultural Conservation Policies

US agricultural conservation policies have been shown to affect farm structure (Ahearn, 2005; Kirwan, 2009). Moreover, farmland prices and rents arguably have both direct and indirect effects on agricultural conservation program participation, and vice versa. For example, Wu and Lin (2010) found a positive correlation between CRP participation and average farmland values in some regions of the United States.

The USDA spends between US\$6 billion and US\$6.5 billion on its six largest conservation programs each year (US Department of Agriculture, 2019). In particular, the CRP is designed to remove environmentally sensitive land from agricultural production in exchange for a yearly rental payment given to the participants for a 10–15 year contract.

Theoretical Approach: Agricultural Land Values

Capitalization Model

According to the capitalization model, agricultural land prices are equal to the present value of the future stream of net benefits to owning land (Duncan, 1979; Castle and Hoch, 1982; Alston, 1986; Burt, 1986; Brueckner, 1990; Nickerson and Zhang, 2014; Kuethe and Oppedahl, 2021; Deaton and Lawley, 2022). In such a case, the present value of land at location i is defined as

$$(1) \qquad V_a(i,t=0) = \int_0^{t^*} R_a(i,t) \times e^{-\rho t}\,dt + \int_{t^*}^{\infty} R_u(i,t,P(t)) \times e^{-\rho t}\,dt - C \times e^{-\rho t^*},$$

where $V_a(i,t=0)$ represents the present, per acre value or price of a parcel of farmland at time $t = 0$.[4] The term $R_a(i,t)$ denotes the present value of a continuous stream of annual net agricultural rents up to time t^*, at which time the farmland is developed for an alternative use with an annual net rental of $R_u(i,t,P(t))$.[5] The term $P(t)$ denotes the population growth path. The parameter ρ represents a constant discount rate or the opportunity cost of owning land. The term C denotes a (constant) conversion cost associated with developing farmland for an alternative use.

If agricultural land is not expected to be converted to alternative uses, then the initial capitalization model can be simplified to

$$(2) \qquad V_a(i,t=0) = \int_0^{\infty} R_a(i,t) \times e^{-\rho t}\,dt.$$

Assuming that the $R_a(i,t)$ is based on annual net returns to agriculture, then the present value of parcel i reduces to a perpetuity formula (Melichar, 1979; Alston, 1986):[6]

$$(3) \qquad V_a(i) = \frac{R_a(i)}{\rho}.$$

This perpetuity formulation can be adjusted to accommodate returns to farmland relative to the discount rate. If future returns to farming grow at a constant rate, $g > 0$, then the above formula can

[4] The capitalization model presented here is very closely related to that of Brueckner (1990).

[5] Alternative uses include development of farmland for urban, residential, or industrial purposes (Deaton and Lawley, 2022).

[6] The current value of net rents to agriculture can be formulated as $R_a(i, t = 0) = \int_0^{\infty} R_a(i, t) \times e^{-\rho t}\,dt$, where ρ is the constant discount rate.

be adjusted as[7]

$$V_a(i) = \frac{R_a(i)}{\rho - g}. \tag{4}$$

The formula in equation (4) allows for the price of farmland to rise (fall) if landowners expect returns to farming to increase (decrease) at the constant growth rate (Deaton and Lawley, 2022).

Similar to equations (3) and (4), Burt (1986) argued that the long-run equilibrium within the market for farmland is given by

$$V_a^*(i) = \alpha \times R_a^*(i), \tag{5}$$

where α is the reciprocal of the constant capitalization rate, $1/\rho$. However, Burt (1986) analyzed the relationship between agricultural rents and land values using a dynamic regression model, in which he treats α as an unknown parameter. Specifically, he analyzed the short-run relationship in equation (5) by specifying the righthand side as a multiplicative, distributed lag of rents as[8]

$$V_a(t) = [\alpha \times R_a(t)^{\beta_0} \times R_a(t-1)^{\beta_1} \times R_a(t-2)^{\beta_2} \ldots] \times u(t). \tag{6}$$

The multiplicative specification implies that market participants consider percentage changes to returns and land prices. Converting equation (6) to natural logs yields

$$\log V_a(t) = \log \alpha + \sum_{j=0}^{\infty} \beta_j \times \log R_a(t-j) + \log u(t). \tag{7}$$

Identification Challenges

Equation (7) offers an elegant representation of the relationship between agricultural land prices and net returns; however, there are several assumptions that belie the equation. Namely, the unknown parameter α encapsulates several important determinants of agricultural land values (e.g., market interest rates, landowner risk preference, and tax policy effects). In this study, we also treat α as an unknown parameter that is controlled for by treating it as unobserved heterogeneous effect. If the unknown parameter α does not control for these factors, then any estimates will potentially suffer from omitted variable bias.

Another weakness is that this specification implicitly assumes that neighboring land values or neighboring rents do not affect a landowner's own value or rents. In reality, interpersonal networks do affect farm learning and decision making (Skaalsveen, Ingram, and Urquhart, 2020), including agricultural land markets. Using a spatial econometric analysis, Patton and McErlean (2003) found that agricultural land prices are not based solely on the inherent characteristics of the land; rather, farmland values reflect the average local price per acre.

To address these challenges, we extend the empirical model in equation (6) in two important ways. First, we utilize a spatial econometric approach (described in the next section), which models own land prices as a spatial-weighted average of neighboring land prices. This extension can be interpreted as a spatiotemporal model; that is, current land values are affected by own past (temporal) prices and neighboring (spatial) price realizations. Intuitively, one could interpret the extension as a proxy for network analysis, which suggests that landowners acquire information about neighboring land values from their peer networks (other regional farmland owners and operators) (Canché, 2018). Second, we utilize a panel data approach that allows us to account for unobserved geographic and temporal heterogeneity, such as the constant market discount rate (represented by the α parameter) discussed above.

[7] The capitalization model in this case can be expressed as $V(i,t) = \int_0^\infty R_a(i,t) \times e^{-(\rho-g)t}\, dt$.

[8] we suppress the county identifier, i, for the rest of this section.

Econometric Model

This study utilizes a dynamic spatial panel data model (Yu, de Jong, and Lee, 2008; Elhorst, Zandberg, and De Haan, 2013; Elhorst, 2014; Yesilyurt and Elhorst, 2017), which can be mathematically expressed as follows:

$$(8) \qquad \boldsymbol{Y}(t) = \tau \times Y(t-1) + \rho \times \boldsymbol{W} \times Y(t) + \eta \times \boldsymbol{W} \times Y(t-1) + \boldsymbol{X}(t)\boldsymbol{\beta} + \boldsymbol{\mu} + \xi(t) + u(t),$$

where $\boldsymbol{Y}(t)$ is an $N \times 1$ vector of observations on the dependent variable for all counties $i = 1, \ldots, N$ for all years in the sample $t = 1, \ldots, T$; $Y(t-1)$ is a one-period time-lagged dependent variable; $\boldsymbol{W}$ is an $N \times N$ row-normalized spatial weight matrix that defines the neighbor relationship between the observations; $W \times Y(t)$ is the endogenous peer effect; $W \times Y(t-1)$ is the time-lagged endogenous peer effect; $\boldsymbol{X}(t)$ is an $N \times K$ matrix of exogenous explanatory variables; $\boldsymbol{\beta}$ is a $K \times 1$ vector of parameters for each explanatory variable; $\boldsymbol{\mu}$ is a vector of spatial fixed effects; $\xi(t)$ is a time fixed effect;[9] and $u(t)$ is the standard *i.i.d.* disturbance term in the model.[10]

The spatial terms in equation (8) offer interpretations that deserve further explanation. The $WY(t)$ term is referred to as the "endogenous peer effect" that controls for how land prices in neighboring counties affect the own-county's prices at time t. In contrast, the term $WY(t-1)$ is the time-lagged endogenous peer effect that controls for how surrounding counties at time $t-1$ affect the own county outcomes today (at time t). It is designed to capture any responses of a county's land market in year t from changes in surrounding counties' land markets in year $t-1$. Additional information regarding the dynamic spatial panel model can be found in Elhorst (2014).

For the sake of space, we omit further discussion of the spatial econometric models. Additional information about the derivation and interpretation of the effects estimates is offered in the online supplement (see www.jareonline.org).

Spatial Weights Matrix

A common misconception within spatial econometrics is that the spatial weights matrix specification will significantly affect a spatial regression model's estimates and any inference drawn from such models. LeSage and Pace (2014) demonstrated that there is little theoretical basis for this misconception. They concluded that the spatial matrix design had little effect on the estimates, provided that an empiricist is using the true partial derivatives (i.e., the marginal effects outlined) and a well-specified spatial regression model.[11] Nonetheless, we provide the estimates based on all three spatial weights matrix designs as a robustness check.

To estimate the spatial econometric models, we prespecified three separate spatial weight matrix designs. The three matrices consisted of (i) a contiguity matrix, (ii) a distance-based matrix, and (iii) a six-nearest-neighbors matrix. The spatial weights are represented by a symmetric square binary matrix, where the values on the main diagonal are operationalized as 0 so that an observational unit cannot be a neighbor to itself. The off-diagonal elements define the spatial relationship between two or more units of observation within a cross section. All of the spatial weights matrices were defined using the "spdep" package in R (Bivand et al., 2022). For the sake of space, we do not define the weight matrices further; however, a more detailed description can be found in the online supplement.

[9] Peer effects in this context refer to neighboring county farmlands.

[10] The *i.i.d.* assumption may be too strong in the context of this particular study, as the agricultural land values and rental rate data are not drawn from a true random sample.

[11] The marginal effects estimates consist of direct, indirect, and total effects (LeSage and Pace, 2009). The direct effects capture the influence of local drivers on the local outcome, whereas the indirect effects capture of influence of neighboring drivers on the local outcome. The indirect effect accounts for spatial spillovers (LeSage and Pace, 2009). The total effect is the sum of the direct and indirect effects. Further definitions of the marginal effects are offered in the online supplement.

Data

We obtained agricultural per acre rental rate data from the USDA National Agricultural Statistical Service's (NASS) Quick Stats service. The rental rate data are based on annual cash rental amounts for nonirrigated cropland (which we refer to as cash rents).[12] Cropland rental rates are measured in US dollars per acre and reported at the county level.[13] The availability of rental rate data forms the basis for the sample. That is, the county-level rental rate data are available from 2008 through 2020, so those were the years of observation chosen for the study. In general, the nonirrigated cash rental rate data are mostly available for counties east of the 100th meridian of the continuous United States (see the map in Figure 2). Counties west of the 100th meridian tend to rely on irrigated agricultural production. Therefore, those counties would likely only have provided irrigated cash rental rates, which are a separate category of rental rate data offered by NASS. For ease of interpretation, we limit the current study to dryland cash rental rates.

Agricultural cash rental rate data are collected annually by NASS to service various USDA programs. Although the cash rental data is very comprehensive, covering over 1,000 counties, there are missing observations in the dataset. Personally identifiable information (PII) restrictions lead NASS to omit a county's cash rental rate observation if the publicly provided data could reveal any proprietary information about a particular agricultural operator (e.g., if there are too few respondents to the survey within a county). Due to these restrictions, the county-level cash rent dataset has been suppressed across some years and counties.

Agricultural land values observations were obtained from CoreLogic, which offers annual real estate sales (transactions) as collected from county tax assessor offices and other sources. The dataset includes a property's location, number of acres, sales price, sales date, assessed value, and property tax information, among other variables. We estimated the per acre sales price by computing the ratio of the individual sale amount relative to the reported acreage in each reported transaction. We limited the analysis to agricultural real estate transactions. It is important to note that we assess all real estate transactions including land with buildings (e.g., farm buildings and residential houses), not just land alone. To compare the individual transaction data to the county-level cash rent data, we aggregated all transactions to the county level by computing the mean value of the county's per acre sales prices. The CoreLogic sales data are missing county-level observations in some years.

The spatial econometric approach in this study does not allow for a significant amount of missing values within the underlying data, without having to drop several observations. This is particularly problematic with the current study as the rental rate observations appear to be missing at random across space and time. To overcome this challenge, we executed a data imputation method utilizing The Comprehensive R Archive Network's "mice" package (van Buuren and Groothuis-Oudshoorn, 2011). The *mice* (multiple imputation by chained equations) package uses a multiple imputation method to train on the observed values of a data matrix and impute the missing values. We experimented with several imputation methods including random forest, predictive mean matching, and regression based. The results of the imputation methods and the evaluation of fit to the data are offered in the online supplement. Based on the fit to the observed data, we chose a predictive mean matching method that imputes a missing data value based on the average of closely associated observed values (van Buuren, 2018).

The data collection and imputation methods offer a cross section of 1,116 counties across 13 years (2008–2020), which led to a final panel data sample of over 14,500 observations. Figure 2 illustrates the counties for which we have cash rent and agricultural real estate transaction data. Dark gray counties in the figure are not due to data suppression by NASS but rather to irrigated versus

[12] Cash rents are a type of land rental contract, similar to residential rental contracts, in which the tenant pays the landowner a fixed monthly rental payment and the farmland renter usually bears all of the production risks. This is in contrast to sharecropping, in which the landowner and farmland renter generally share the production risks (Bigelow, Borchers, and Hubbs, 2016).

[13] Going forward, we will refer to the county-level rents as agricultural rents, but the observed rents are technically cropland rents as we do not include data on rangeland or irrigated cropland rents, which NASS reports separately.

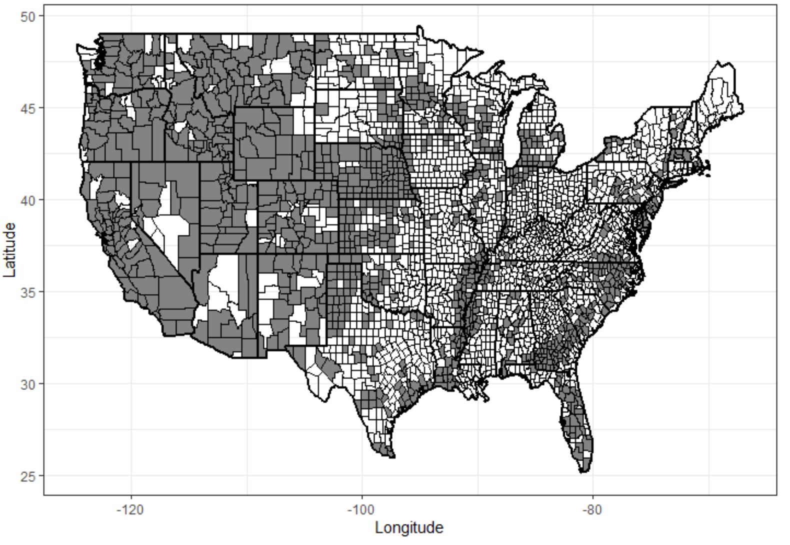

Figure 2. County-Level Sample of Rental and Agricultural Land Values, 2008–2020

Notes: Dark gray areas indicate missing or sparse observations. These regions rely heavily upon irrigation, so reported rental rates would be for irrigated croplands, which are outside of the scope of the current study. White areas indicate observed counties, which are largely nonirrigated croplands, and therefore within the scope of the study.
Sources: Map created by authors based on the availability of NASS's county-level cash rents data.

nonirrigated croplands. We only analyze nonirrigated cropland rental rates. The divide between the observed (white) counties and omitted (dark gray) counties basically falls along the 100th meridian (the −100 longitude measure on the map), which forms the boundary between the humid Eastern United States and the arid Western plains. Many agricultural regions west of the 100th meridian rely upon irrigation, so only irrigated cropland rental rates would be reported in such regions. Examining both irrigated and nonirrigated rental rates would add an unnecessary level of complexity to the current empirical approach.

To convert nominal cash rents, agricultural land values, and operational returns to real dollars, we utilized the US Bureau of Economic Analysis's gross domestic product implicit price deflator (Federal Reserve Bank of St. Louis, 2022).

Cash receipts are defined as total cash receipts from marketing. The US Bureau of Economic Analysis (2019) estimates farm income by collecting initial observations from Census of Agriculture data. Since the agricultural census is only available every 5 years, the BEA estimates intercensal (annual) cash receipts amounts by using other regional economic account data.

Farming practices and crop choices vary regionally according to climate and meteorological conditions, so we also collected data on (county-level) average temperatures and precipitation using data from the US National Oceanic and Atmospheric Administration (2022) National Centers for Environmental Information. Finally, we collected total and CRP crop acreage data from the USDA's (US Department of Agriculture, 2022a). The Farm Service Agency reports annual county-level totals of agricultural acres enrolled in the CRP.

Table 1 offers summary statistics. The mean cash rental rate was approximately \$160/acre, whereas the average agricultural land value was nearly \$6,400/acre over the period of observation. The average per acre price is slightly higher than the USDA's calculation of \$4,100/acre, per the June Area Survey (US Department of Agriculture, 2022b). However, our estimates are based on transaction data instead of survey-elicited values, and the estimated median value of \$4,566/acre is close to the June Area Survey average. The CoreLogic data are recorded by the local tax assessor's office, and the reported values include land as well as improvements thereon (e.g., buildings and other assets).

Table 1. Summary Statistics for Data Sample, 2008–2020

Variable	Units	Min.	Q1	Median	Mean	Q3	Max.
Market determinants							
Cash rent	$US/acre	16.00	86.00	130.00	158.00	193.00	2,610.00
CoreLogic sales	$US/acre (thousands)	0.00	2.90	4.57	6.44	8.14	73.02
County-level socioeconomics							
Population	Thousands (entire county)	0.44	9.50	25.29	124.69	77.58	10,094.87
Median age	Years (entire county)	21.50	36.50	39.90	40.06	43.50	68.50
Farm and geographic covariates							
Average temperatures	Degrees Fahrenheit	33.94	47.63	53.69	54.70	61.86	78.38
Average precipitation	Inches	0.10	1.66	2.76	2.90	3.96	10.69
CRP acreage	Thousands	0.00	0.87	4.34	15.35	14.41	276.81
Cash receipts	$US (thousands)	2.89	53.45	109.92	200.78	206.67	6,750.88

Notes: Q1 and Q3 denote the median estimated value within the first and third quantile of the distribution. Cash rents, CoreLogic sales price, rent, and farm income are provided in real dollars (as adjusted by the GDP deflator) The final panel dataset is full balanced, with $N = 14,508$.

We hypothesize that cash rental rates—as a proxy for farm returns—are positively associated with cropland land values. This hypothesis is based, in part, on the past work of Kirwan (2009), who argued that agricultural rental rates respond primarily to innovations in expected farm returns in the current period. As further evidence of the relationship, we conducted a series of auxiliary regressions between county-level cash rental rates and county-level farm returns using the Census of Agriculture data. Similar to Kirwan (2009) and Gutierrez, Westerlun, and Erickson (2007), our estimates suggest a linear relationship between rental rates and returns. These estimates are provided in the online supplement.

Brueckner (1990) argued that the county-level population growth path is an important determinant of whether a landowner chooses to develop agricultural land into other residential, urban, or industrial purposes. Further, we collected data (from the US Census Bureau) on the median age of the entire population within the county, as the average age of US agricultural producers has been steadily rising over the past few decades.

Results

The baseline of comparison is two nonspatial fixed effects estimates, which are offered in columns 1 and 2 of Table 2. Column 1 presents pooled ordinary least squares (OLS) estimates based on the raw or nonimputed data, and column 2 presents fixed effects estimates based on the imputed data. As discussed in the previous section, the cash rental rate data (collected by NASS) and the land sales transaction data (collected by CoreLogic) contained missing values over space and time. The missing variables can cause biased and inefficient parameter estimates (Greene, 2011). Due to the missing data, the regression results in column 1 are based on fewer than 2,800 observations, and the regressions using the imputed data, in columns 2–5, contain over 14,500 observations.

Columns 3–5 in Table 2 offer the comparative spatial fixed effects estimates: column 3 contains a spatial lag of the dependent variable; column 4 contains a spatial lag of the error term; and column 5 contains a spatial lag of both the dependent variable and the error term. The spatially lagged dependent model assumes that spatial connectivities between neighboring markets directly affect local agricultural land values (and vice versa). The specification for the spatially lagged error term assumes that there is spatial dependence within shocks to the market values. The third specification assumes spacial connectivities within prices as well as market shocks.[14]

Arguably because of the missing data problem, the coefficient estimates for many of the variables in column 1 are larger in absolute magnitude than the estimates from the other models. In column 1, the coefficient on the cash rental rate implies that a 10% increase in rates is positively correlated with an approximate 1.3% increase in agricultural land values. The estimated coefficients on cash rents in columns 2–5 are relatively consistent with those in column 1. As indicated above, the rental rate data are a proxy for agricultural returns, so the coefficient implies that an increase in per acre returns is positively associated with an increase in land values. All of the spatial econometric models assume a contiguity spatial weights matrix (i.e., if two counties share a common border, then the weights matrix contains a binary assignment of unity representing a shared border).[15]

In comparison, the estimated coefficient on cash receipts—the alternative measure of returns—is positive, but the estimate is not statistically different from 0 for any of the nonspatial or spatial estimates. The lack of significance is arguably due to the fact that the cash receipts from marketing variable is estimated and not necessarily based on actual annual returns.

The estimated coefficient on CRP acreage in Table 2 implies a significant, inverse relationship with average farmland values. The coefficient estimate, based on the nonimputed data in column 1, is much larger in absolute magnitude than the other model estimates. Again though, the estimate is column 1 is arguably biased due to the missing data problem. The estimated effect of CRP acreage is quite small for the models in columns 2–5. The estimate in column 2 suggests that a 10% increase in county-level CRP acres is associated with a 0.08% decrease in average farmland values. At the sample average of farmland values, approximately $6,444/acre, the 0.08% diminution in value would equate to about $5/acre.

The negative estimate on CRP acreage may seem counterintuitive, as Wu and Lin (2010) found a positive relationship. However, the negative relationship is consistent with the first-order condition of the land capitalization model. To explain, consider the insights of Taylor et al. (2020), who claimed that the length of CRP contracts (10–15 years) and restrictions on land use introduce an opportunity cost equal to the highest alternative use of land. Moreover, farmers must incur significant costs to convert CRP land back into crop production (Roberts and Lubowski, 2007). The opportunity costs are the foregone returns from crop production or other nonagricultural development options. Taylor et al. argued that if commodity prices are increasing during a contract period, then nearby farmland available for crop production is likely to experience increasing rental rates.

As motivated in the theory section above, much of the past literature has implicitly assumed that own agricultural land values are independent of neighboring values. The estimated spatial autocorrelation coefficients in columns 3 and 4 of Table 2 suggest otherwise. These estimates imply that farmland values are influenced by neighboring county values (Hardie, Narayan, and Gardner, 2001; Huang et al., 2006).[16] Ignoring neighboring effects or spatial dependencies may affect the efficiency and consistency of estimates (Kim, Phipps, and Anselin, 2003).

The regression results in Table 2 are highly suggestive of spatial correlation between neighboring agricultural land values; however, these specifications omit potentially important temporal dynamics. In Table 3 we include the estimates for the dynamic spatial panel data model.[17] A caveat with the dynamic spatial panel data model is that the coefficient estimates may be biased if the initial cross section of observations is treated as exogenous. Parent and LeSage (2011) developed a Bayesian hierarchical model that controls for random effects; their model includes a spatial filter that allows the researcher to specify whether the initial cross section of observations is endogenous

[14] Such a specification could be represented as $y = \rho \times Wy + X\beta + \varepsilon$, $\varepsilon = \phi \times u$, where ϕ is the spatial autocorrelation coefficient on the error term.

[15] We experimented with other spatial weights matrices. The estimates—provided in the online supplement—are nearly identical, with very slight variation in some of the coefficient estimates.

[16] The spatial panel data econometric models were estimated using the "splm" package in R (Millo, Piras, and Bivand, 2022).

[17] The dynamic spatial panel data models were estimated using the R package "SDPDmod" (Simonovska, 2022).

Table 2. Nonspatial and Spatial Fixed Effects Estimates

Variable	Nonimputed (*N* = 2,791) Pooled OLS 1	Imputed Data (*N* = 14,508) FE Estimator 2	Spatial Lag 3	Spatial Error 4	Sp. Lag & Error 5
log(Cash rent)	0.1262***	0.1267***	0.1262***	0.1264***	0.1261***
	(0.0472)	(0.0127)	(0.0110)	(0.0110)	(0.0110)
log(Cash receipts)	0.0494	0.0325	0.0325	0.0341	0.0324
	(0.0293)	(0.0395)	(0.0318)	(0.0321)	(0.0318)
log(CRP acres)	−0.0454***	−0.0083*	−0.0083**	−0.0084**	−0.0083**
	(0.0103)	(0.0037)	(0.0031)	(0.0031)	(0.0031)
log(Population)	0.1819***	0.0447	0.0394	0.0359	0.0398
	(0.0211)	(0.1551)	(0.1279)	(0.1300)	(0.1277)
Median age	−0.0150**	−0.0004	−0.0004	−0.0003	−0.0004
	(0.0056)	(0.0062)	(0.0050)	(0.0051)	(0.0050)
Avg. temperature	−0.0204***	0.0064	0.0062	0.0064	0.0061
	(0.0029)	(0.0060)	(0.0054)	(0.0056)	(0.0055)
Avg. precipitation	0.0167	0.0156	0.0149	0.0153	0.0148
	(0.0205)	(0.0088)	(0.0090)	(0.0093)	(0.0092)
Spatial autocorrelation coefficients					
Spatial lag			0.0370***		0.0414
			(0.0111)		(0.1334)
Spatially lagged error				0.0366**	−0.0045
				(0.0111)	(0.1355)
R^2	0.281	0.461	0.461	0.461	0.461
AIC	5,266	27,474	25,204	25,220	25,205
BIC	5,313	36,080	25,265	25,281	25,273

Notes: The dependent variable is the log of the average per acre sales price. Agricultural land sale price, rental rate, and farm income data were converted to real values using the GDP deflator. Single, double, and triple asterisks (*, **, ***) indicate significance at the 1%, 5%, and 10% level, respectively. Values in parentheses are estimated standard errors. The regression models in columns 3–5 are based on a contiguity spatial matrix design. Estimates based on the distance and nearest neighbor matrix designs are available in the online supplement. CRP denotes the Conservation Reserve Program, AIC indicates the Akaike information criterion, and BIC indicates the Bayesian informationn criterion. All regressions control for spatial and year fixed effects.

or exogenous. Unfortunately, the framework presented here does not allow for such a spatial filter specification, so the current approach assumes that the initial cross section is exogenous.

As displayed in Table 3, the estimated coefficient on cash rents remains positive and highly statistically significant with each specification. The coefficients are slightly smaller relative to the spatial model estimates in Table 2. The coefficient estimate in column 1 suggests that a 10% increase in agricultural returns (proxied by cash rents) is associated with an approximate 1.2% increase in agricultural land values.

The spatial autocorrelation coefficient estimates are positive and significant, suggesting once again that neighboring land values are an important determinant of own-county values. Unfortunately, there is no real direct interpretation of these coefficients other than the estimated sign and consistency in statistical significance across specifications. However, we can decompose the estimates (described in the online supplement) into direct, indirect, and total effects, which are provided in Table 4. Recall that direct effects capture the effect of the local explanatory variable on the local dependent variable, whereas the indirect effects capture the effect of the neighboring explanatory variable on the local dependent variable. Thus, the indirect effects arguably account for spatial spillovers. The total effect is the sum of the direct and indirect effects.

The results in Table 4 arguably reflect a more accurate estimate of the marginal effects as the estimates capture the temporal dynamics and spatial connectivities within the agricultural land markets. The effects estimates are divided into short-run impacts (upper portion of the table) and long-run impacts (lower portion of the table). The results in Table 4 are based on the contiguity spatial weights matrix. As a robustness check, we also experimented with the previous mentioned distance- and nearest neighbors-based weights matrices. The results, which are nearly identical, are provided in the online supplement.[18]

According to the results in Table 4, the only significant determinants of per acre cropland values are per acre returns (proxied by cash rental rates) and CRP acres. As before, the coefficient estimates on cash rents remain positive and highly statistically significant. The direct, short-run impact implies that a 10% increase in rents is associated with an own-county increase in average land values of about 1.2%. Based on the sample mean of land values, the estimate suggests that a 1% increase in cash rental rates (returns) is associated with a nearly \$8 increase in per acre land values. The short-run indirect estimate suggests that a 10% increase in neighboring rental rates (returns) is associated with a 0.03% increase in average land values. The indirect effect estimates for CRP acreage are marginally significant at the 10% level. The estimated long-run impact of returns (cash rents) is larger in magnitude than the short-run estimate. For example, the estimated long-run total impact on cash rents is about 0.14, which is approximately a \$9 per acre average increase in land values.

The estimated coefficients on CRP acreage in Table 4 suggest that additional CRP participation is negatively correlated with agricultural land values. The short-run direct impact estimate implies that a 10% increase of within-county CRP participation is associated with an approximate 0.1% decrease in land values. The short-run indirect effect is not significant, suggesting that neighboring county CRP participation is not correlated with local farm values. The negative estimates on CRP acres in the long run are only slightly larger in magnitude (in absolute terms). For example, the total long-run impact estimate suggests that a 10% increase in CRP participation is associated with a \$5/acre diminution in land values.

Forecasts of Nonspatial and Spatial Econometric Models

In the previous subsection, we implicitly argued that the dynamic spatial panel data model should provide superior estimates over the nonspatial and static spatial fixed effects models. However, the estimates for the dynamic spatial panel data model in Tables 3 and 4 are based on in-sample observations. As further evidence of the accuracy of the dynamic spatial panel data method, we tested the forecasting ability of each of the models.

In order to carry out the forecasts, we defined the first 10 years (2008–2017) of observations as the training set and the last 3 years (2018–2020) as the test set. In other words, we estimated the dynamic spatial panel data model using the training data and then predicted the per acre farmland sales prices for the last 3 years of data. One could interpret the in sample as the training dataset and the out of sample as the test dataset. To calculate the forecasted per acre prices, we multiplied the coefficient estimates from the training-data regressions by the explanatory variables in the test dataset.

After estimating the predicted values, we calculated five separate error forecasting metrics, including the mean absolute error (MAE), root mean squared error (RMSE), normalized root mean squared error (NRMSE), weighted average percentage error (WAPE), and mean absolute percentage error (MAPE). Each of these metrics evaluates the forecasting errors (the out-of-sample equivalent to the in-sample residuals) based on different assumptions of the data-generating process (Allen and Fildes, 2001). Further definitions and explanations of the five forecasting metrics are provided in the online supplement.

Table 5 reports the forecasting error metrics. In general, the lower the estimated metric, the better the relative model forecast (i.e., the smaller the forecasting error). As can be gleaned in the table, the nonspatial fixed effects model provides the least accurate out-of-sample forecasts, whereas the spatial autoregressive model provides the second-best forecasts. As implied in the previous

[18] The short- and long-run marginal effects estimates differ only marginally.

Table 3. Dynamic Spatial Panel Data Estimation with Fixed Effects

Spatial Weights	Contiguity 1	Distance 2	Nearest Neighbor 3
log(Cash rent)	0.1172***	0.1164***	0.1168***
	(0.0120)	(0.0120)	(0.0120)
log(Sale price $(t-1)$)	0.0413***	0.0415***	0.0415***
	(0.0089)	(0.0089)	(0.0089)
W*log(Sales price $(t-1)$)	0.0610***	0.0291	0.0488*
	(0.0169)	(0.0179)	(0.0211)
log(Cash receipts)	0.0248	0.0253	0.0238
	(0.0354)	(0.0354)	(0.0354)
log(CRP acreage)	−0.0065*	−0.0064	−0.0065
	(0.0032)	(0.0032)	(0.0033)
log(Population)	0.0991	0.1034	0.1005
	(0.1447)	(0.1447)	(0.1448)
Average temperature	0.0058	0.0053	0.0055
	(0.0057)	(0.0057)	(0.0057)
Average precipitation	0.0179	0.0173	0.0176
	(0.0095)	(0.0095)	(0.0095)
Median age	−0.0010	−0.0012	−0.0013
	(0.0060)	(0.0060)	(0.0060)
Spatial autocorrelation coefficient			
Spatially lagged dependent	0.0270*	0.0568***	0.0457***
	(0.0116)	(0.0131)	(0.0156)
R^2	0.465	0.464	0.465
AIC	22,938	22,932	22,940
BIC	23,006	23,000	23,008

Notes: The dependent variable is the log of agricultural per acre sale price. Agricultural land values, rental rates, and farm income data were converted to real values using the GDP deflator. Single, double, and triple asterisks (*, **, ***) indicate significance at the 1%, 5%, and 10% level, respectively. Values in parentheses are estimated standard errors. Other variables not included are average precipitation. CRP denotes the Conservation Reserve Program, AIC indicates the Akaike information criterion, and BIC indicates the Bayesian informationn criterion. . The three column headings indicate the type of spatial weights matrix used for the estimation of the models. Column 1 denotes a contiguity-based weights matrix, column 2 indicates a distance-based weights matrix, and column 3 indicates a nearest neighbor-based weights matrix.

subsection, the dynamic spatial panel data model provides superior out-of-sample forecasts relative to all the other models.

Discussion

In this study, we analyzed the relationship between agricultural returns (proxied by farmland rental rates) and agricultural land values at the county level from 2008 to 2020. Using conventional (i.e., nonspatial) fixed effects estimation, we found that cash rental rates and CRP participation were significant determinants of average per acre agricultural land values.

As agricultural land values are likely affected by spatially interconnected market activities, we posited that a spatial econometric fixed effects estimator would arguably better represent the farmland price mechanism. As such, we utilized several specifications of a spatial econometric fixed effects estimator. The spatial autocorrelation coefficients within each of these models was found

Table 4. Direct, Indirect, and Total Impact Estimates Based on Dynamic Spatial Panel Data Models

	Direct 1	Indirect 2	Total 3
Short-run impacts			
log(Cash rents)	0.1182***	0.0031	0.1212***
	(0.0121)	(0.0016)	(0.0125)
log(Cash receipts)	0.0228	0.0006	0.0234
	(0.0364)	(0.0011)	(0.0373)
log(CRP acres)	−0.0065*	−0.0002	−0.0067*
	(0.0032)	(0.0001)	(0.0033)
log(Population)	0.1142	0.0029	0.1170
	(0.1367)	(0.0044)	(0.1403)
Average temperature	0.0061	0.0002	0.0063
	(0.0053)	(0.0002)	(0.0054)
Average precipitation	0.0181	0.0005	0.0186
	(0.0094)	(0.0004)	(0.0096)
Median age	0.0020	0.0001	0.0020
	(0.0064)	(0.0002)	(0.0065)
Long-run impacts			
log(Cash rents)	0.1236***	0.0115***	0.1351***
	(0.0127)	(0.0031)	(0.0141)
log(Cash receipts)	0.0238	0.0022	0.0261
	(0.0380)	(0.0038)	(0.0417)
log(CRP acres)	−0.0068*	−0.0006	−0.0074*
	(0.0034)	(0.0004)	(0.0037)
log(Population)	0.1194	0.0109	0.1303
	(0.1428)	(0.0140)	(0.1561)
Average temperature	0.0064	0.0006	0.0070
	(0.0055)	(0.0006)	(0.0060)
Average precipitation	0.0190	0.0017	0.0207
	(0.0098)	(0.0010)	(0.0107)
Median age	0.0021	0.0002	0.0023
	(0.0066)	(0.0006)	(0.0072)

Notes: The dependent variable is the log of per acre agricultural land values. All impact estimates are based on the dynamic spatial panel data estimator with the contiguity spatial weights matrix. Agricultural land values, rental rates, and farm income data were converted to real values using the GDP deflator. Single, double, and triple asterisks (*, **, ***) indicate significance at the 1%, 5%, and 10% level, respectively. Values in parentheses are estimated standard errors. CRP denotes the Conservation Reserve Program.

to be statistically significant, suggesting that neighboring agricultural land values are positively correlated with own local land values (and vice versa).

The dynamic spatial panel data models offered a more nuanced interpretation of the regression results, and we contend that the derived marginal effects estimates better reflect the spatial connectivities and temporal dynamics within agricultural land markets. Using the coefficient estimates, we decomposed the effects into direct, indirect, and total impacts in the short and long run. The impacts suggested that cash rental rates—a proxy for agricultural returns—are positively associated with own and neighboring farmland values in the short and long run. Moreover, we

Table 5. Forecasting Error Metrics

Model	MAE	RMSE	NRMSE	WAPE	MAPE
Nonspatial fixed effects	8.5091	8.5825	13.9312	0.9235	0.9305
Spatial autoregressive	0.7113	0.9245	0.1028	0.0779	0.0796
Spatial error model	0.7137	0.9256	0.1010	0.0782	0.0782
Spatial autocorrelation	1.1427	1.3672	0.1671	0.1252	0.1234
Dynamic spatial model	0.2914	0.5399	0.0593	0.0320	0.0347

Notes: Table 5 displays the metrics used to evaluate the forecasting error of each model. MAE is mean absolute error, RMSE is root mean squared error, NRMSE is normalized root mean squared error, WAPE is weighted average percentage error and MAPE is mean absolute percentage error. The online supplement provides definitions for each metric. All of the models were estimated using two-way fixed effects. To estimate the out-of-sample forecasts, we ran the regressions on a 2008–2017 partition of the data then used the parameter estimates to calculate the fitted values for the out-of-sample period (2018–2010). The spatial econometric models were estimated using the contiguity-based spatial weights matrix. Similar metrics were calculated for the spatial econometric models with the distance- and nearest neighbor-based weight matrices—these metrics are not provided but are available upon request. The results were the same despite the specified weight measure (i.e., the dynamic spatial panel data model provides the lowest forecasting errors).

demonstrated that the dynamic spatial panel data model provided the most accurate out-of-sample forecasts of per acre agricultural land transactions in 2018–2020.

Domestic policymakers are conscious that compensating operators with distortionary rents would negatively affect local agricultural markets. There has been little past research investigating the relationship between CRP participation and agricultural land values. Based on the estimates within the dynamic spatial panel data approach), we found that CRP-enrolled land in own and neighboring counties may have a small negative effect on per acre farmland values. Our findings were different from those of Wu and Lin (2010), who found a positive relationship between CRP participation and farmland values. We argue that the difference with our estimates is possibly due to the opportunity costs (e.g., foregone returns associated with rising commodity prices over time) of having a fixed-payment contract for a period of 10–15 years. Nevertheless, the Farm Service Agency could potentially use these methods to inform the Conservation Reserve Program in the future.

This study was subject to limitations. Namely, the cash rental rate data for nonirrigated cropland and CoreLogic real estate transaction data contained missing observations. As mentioned above, the cross-sectional survey data (which is observed annually) is fairly comprehensive, but the National Agricultural Statistics Service omits reporting cash rental rates in some counties due to potentially exposing personally identifiable information. To overcome this problem, we had to impute rental rate and cropland sales prices. Since the rental and sales data appear to be missing at random, the multiple imputation by chained equations method should provide sufficient estimates. However, future research may consider other sources of agricultural returns data or other methods of imputation.

[First submitted October 2022; accepted for publication April 2023.]

References

Ahearn, M. 2005. "Effects of Differing Farm Policies on Farm Structure and Dynamics." *American Journal of Agricultural Economics* 87(5):1182–1189. doi: 10.1111/j.1467-8276.2005.00805.x.

Allen, P. G., and R. Fildes. 2001. "Econometric Forecasting." In J. Armstrong, ed., *Principles of Forecasting: A Handbook for Researchers and Practitioners, Technological Forecasting and Social Change*, Vol. 69(3). New York, NY: Springer, 313–316. doi: 10.1016/S0040-1625(02)00180-4.

Alston, J. M. 1986. "An Analysis of Growth of U.S. Farmland Prices, 1963–1982." *American Journal of Agricultural Economics* 68(1):1–9. doi: 10.1016/0169-2070(94)90051-5.

Benirschka, M., and J. K. Binkley. 1994. "Land Price Volatility in a Geographically Dispersed Market." *American Journal of Agricultural Economics* 76(2):185–195. doi: 10.2307/1243620.

Bigelow, D., A. Borchers, and T. Hubbs. 2016. *U.S. Farmland Ownership, Tenure, and Transfer*. Economic Information Bulletin EIB-161. Washington, DC: USDA Economic Research Service. doi: 10.22004/ag.econ.262138.

Bigelow, D. P., and T. Hubbs. 2016. "Land Acquisition and Transfer in U.S. Agriculture." *Amber Waves.* doi: 10.22004/ag.econ.245672.

Bigelow, D. P., J. Ifft, and T. Kuethe. 2020. "Following the Market? Hedonic Farmland Valuation Using Sales Prices versus Self-Reported Values." *Land Economics* 96(3):418–440. doi: 10.3368/le.96.3.418.

Bivand, R., M. Altman, L. Anselin, R. Assunçaõ, O. Berke, G. Blanchet, M. Carvalho, B. Christensen, Y. Chun, C. Dormann, S. Dray, and G. Piras. 2022. "spdep: Spatial Dependence: Weighting Schemes, Statistics." R v1.2-5.

Brueckner, J. K. 1990. "Growth Controls and Land Values in an Open City." *Land Economics* 66(3): 237–248. doi: 10.2307/3146726.

Burns, C., N. Key, S. Tulman, A. Borchers, and J. Weber. 2018. *Farmland Values, Land Ownership, and Returns to Farmland, 2000–2016*. Economic Research Report ERR-245. Washington, DC: USDA Economic Research Service. doi: 10.22004/ag.econ.276249.

Burt, O. 1986. "Econometric Modeling of the Capitalization Formula for Farmland Prices." *American Journal of Agricultural Economics* 68(1):10–26. doi: 10.2307/1241645.

Callahan, S. 2020. "Farmland Value." Washington, DC: USDA Economic Research Service. Available online at https://www.ers.usda.gov/topics/farm-economy/land-use-land-value-tenure/farmland-value/.

Canché, M. S. G. 2018. "Spatial Econometrics and Network Analysis as Means to Assess the Assumption of Independence in Higher Education Research." *New Directions for Institutional Research* 180:11–38. doi: 10.1002/ir.20284.

Castle, E., and I. Hoch. 1982. "Farm Real Estate Price Components, 1920–78." *American Journal of Agricultural Economics* 64:8–18. doi: 10.2307/1241167.

Deaton, B., and C. Lawley. 2022. "A Survey of Literature Examining Farmland Prices: A Canadian Focus." *Canadian Journal of Agricultural Economics* 70(2):95–121. doi: 10.1111/cjag.12307.

Dillard, J. G., T. H. Kuethe, C. Dobbins, M. Boehlje, and R. J. G. M. Florax. 2013. "The Impacts of the Tax-Deferred Exchange Provision on Farm Real Estate Values." *Land Economics* 89(3): 479–489. doi: 10.3368/le.89.3.479.

Duncan, M. 1979. *Farm Feal Estate Values: What's Happening and Why*. Kansas City, MO: Federal Reserve Bank of Kansas City.

Elhorst, J. P. 2014. "Dynamic Spatial Panels: Models, Methods and Inferences." In *Spatial Econometrics,* Springer, 95–119. doi: 10.1007/s10109-011-0158-4.

Elhorst, P., E. Zandberg, and J. De Haan. 2013. "The Impact of Interaction Effects among Neighbouring Countries on Financial Liberalization and Reform: a Dynamic Spatial Panel Data Approach." *Spatial Economic Analysis* 8(3):293–313. doi: 10.1080/17421772.2012.760136.

Federal Reserve Bank of St. Louis. 2022. "Gross Domestic Product: Implicit Price Deflator." Available online at https://fred.stlouisfed.org/series/GDPDEF#0 [Accessed August 2022].

Greene, W. H. 2011. *Econometric Analysis*, 7th ed. New York, NY: Pearson.

Gutierrez, L., J. Westerlun, and K. Erickson. 2007. "Farmland Prices, Structural Breaks and Panel Data." *European Review of Agricultural Economics* 34(2):161–179. doi: 10.1093/erae/jbm018.

Hardie, I. W., T. A. Narayan, and B. L. Gardner. 2001. "The Joint Influence of Agricultural and Non-Farm Factors on Real Estate Values: An Application to the Mid-Atlantic Region." *American Journal of Agricultural Economics* 83(1):120–132. doi: 10.1111/0002-9092.00141.

Huang, H., G. Y. Miller, B. J. Sherrick, and M. I. Gómez. 2006. "Factors Influencing Illinois Farmland Values." *American Journal of Agricultural Economics* 88(2):458–470. doi: 10.1111/j.1467-8276.2006.00871.x.

Johnson, B. 2010. “A Thin Real Estate Market Becomes Even Leaner.” *Cornhusker Economics* 492. Available online at https://digitalcommons.unl.edu/agecon_cornhusker/492/ [Accessed June 2022].

Kim, C. W., T. T. Phipps, and L. Anselin. 2003. “Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach.” *Journal of Environmental Economics and Management* 45(1):24–39. doi: 10.1016/S0095-0696(02)00013-X.

Kirwan, B. E. 2009. “The Incidence of U.S. Agricultural Subsidies on Farmland Rental Rates.” *Journal of Political Economy* 117(1):138–164. doi: 10.1086/598688.

Kuethe, T. H., and J. Ifft. 2013. “The Information Content of Farmland Value Surveys.” *Agricultural Finance Review* 73(1):45–57. doi: 10.1108/00021461311321302.

Kuethe, T. H., and D. Oppedahl. 2021. “Agricultural Bankers’ Farmland Price Expectations.” *European Review of Agricultural Economics* 48(1):42–59. doi: 10.1093/erae/jbaa017.

LeSage, J., and R. K. Pace. 2009. *Introduction to Spatial Econometrics*. Chapman and Hall/CRC.

LeSage, J. P., and R. K. Pace. 2014. “The Biggest Myth in Spatial Econometrics.” *Econometrics* 2(4):217–249. doi: 10.3390/econometrics2040217.

Melichar, E. 1979. “Capital Gains versus Current Income in the Farming Sector.” *American Journal of Agricultural Economics* 61(5):1058–1092. doi: 10.2307/3180381.

Millo, G., G. Piras, and R. Bivand. 2022. “‘splm’: Econometric Models for Spatial Panel Data.” R package v1.6-2.

Nickerson, C. J., M. Morehart, T. Kuethe, J. Beckman, J. Ifft, and R. Williams. 2012. *Trends in U.S. Farmland Values and Ownership*. Economic Information Bulletin EIB-92. Washington, DC: USDA Economic Research Service.

Nickerson, C. J., and W. Zhang. 2014. “Modeling the Determinants of Farmland Values in the United States.” In J. M. Duke and J. Wu, eds., *The Oxford Handbook of Land Economics,* Oxford, UK: Oxford University Press, 111–138.

Nolte, C. 2020. “High-Resolution Land Value Maps Reveal Underestimation of Conservation Costs in the United States.” *Proceedings of the National Academy of Sciences* 117(47):29577–29583. doi: 10.1073/pnas.2012865117.

Parent, O., and J. P. LeSage. 2011. “A SpaceâĂŞTime Filter for Panel Data Models Containing Random Effects.” *Computational Statistics & Data Analysis* 55(1):475–490. doi: 10.1016/j.csda. 2010.05.016.

Patton, M., and S. McErlean. 2003. “Spatial Effects within the Agricultural Land Market in Northern Ireland.” *Journal of Agricultural Economics* 54(1):35–54. doi: 10.1111/j.1477-9552. 2003.tb00047.x.

Plantinga, A. J., R. N. Lubowski, and R. N. Stavins. 2002. “The Effects of Potential Land Development on Agricultural Land Prices.” *Journal of Urban Economics* 52:561–581. doi: 10.1016/S0094-1190(02)00503-X.

Roberts, M. J., and R. N. Lubowski. 2007. “Enduring Impacts of Land Retirement Policies: Evidence from the Conservation Reserve Program.” *Land Economics* 83(4):516âĂŞ538. doi: 10.3368/le.83.4.516.

Robison, L. J., and S. R. Koenig. 1992. “Market Value versus Agricultural Use Value of Farmland.” In M. C. Ahearn and U. Vasavada, eds., *Costs and Returns for Agricultural Commodities,* New York, NY: Routledge, 207–228. doi: 10.1201/9780429036385-25.

Simonovska, R. 2022. “SDPDmod: Spatial Dynamic Panel Data Modeling.” R package v0.0.0.9000.

Skaalsveen, K., J. Ingram, and J. Urquhart. 2020. “The Role of Farmers’ Social Networks in the Implementation of No-Till Farming Practices.” *Agricultural Systems* 181:1–14. doi: 10.1016/j.agsy.2020.102824.

Soule, M. J., A. Tegene, and K. D. Wiebe. 2000. “Land Tenure and the Adoption of Conservation Practices.” *American Journal of Agricultural Economics* 82(2):993–1005. doi: 10.1111/0002-9092.00097.

Taylor, M. R., N. P. Hendricks, G. S. Sampson, and D. Garr. 2020. "The Opportunity Cost of the Conservation Reserve Program: A Kansas Land Example." *Applied Economic Perspectives and Policy* 43(2):849–865. doi: 10.1002/aepp.13040.

US Bureau of Economic Analysis. 2019. "A Primer on Local Area Gross Domestic Product Methodology." Available online at https://apps.bea.gov/scb/2020/03-march/0320-county-level-gdp.htm [Accessed February 2022].

US Department of Agriculture. 2019. "Conservation Programs: Conservation Spending Seeks to Improve Environmental Performance in Agriculture." Washington, DC: USDA Economic Research Service. Available online at https://www.ers.usda.gov/topics/natural-resources-environment/conservation-programs/ [Accessed November 2021].

———. 2022a. "Crop Acreage Data." Washington, DC: USDA Farm Service Agency. Available online at https://www.fsa.usda.gov/news-room/efoia/electronic-reading-room/frequently-requested-information/crop-acreage-data/index [Accessed February 2022].

———. 2022b. "Surveys: June Area." Washington, DC: USDA National Agricultural Statistics Service. Available online at https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/June_Area [Accessed August 2022].

US National Oceanic and Atmospheric Administration. 2022. "NOAA Offers Climate Data for Counties." Available online at https://www.ncei.noaa.gov/news/noaa-offers-climate-data-counties [Accessed Februrary 2022].

van Buuren, S. 2018. *Flexible Imputation of Missing Data*, 2nd ed. New York, NY: CRC.

van Buuren, S., and K. Groothuis-Oudshoorn. 2011. "mice: Multivariate IMputation by Chained Equations in R." *Journal of Statistical Software* 45(3):1–67. doi: 10.18637/jss.v045.i03.

Vyhnalek, A., J. Jansen, J. Groskopf, and A. Duerfeldt. 2020. "Frequently Asked Questions: Farmland Leases." University of Nebraska-Lincoln Department of Agricultural Economics. Available online at https://agecon.unl.edu/documents/land/farm-lease-faq2020.pdf [Accessed December 2022].

Wentland, S. A., Z. H. Ancona, K. J. Bagstad, J. Boyd, J. L. Hass, M. Gindelsky, and J. G. Moulton. 2020. "Accounting for Land in the United States: Integrating Physical Land Cover, Land Use, and Monetary Valuation." *Ecosystem Services* 46:101178. doi: 10.1016/j.ecoser.2020.101178.

Wu, J.-J., and H. Lin. 2010. "The Effect of the Conservation Reserve Program on Land Values." *Land Economics* 86(1):1–21. doi: 10.3368/le.86.1.1.

Yesilyurt, M. E., and J. P. Elhorst. 2017. "Impacts of Neighboring Countries on Military Expenditures: A Dynamic Spatial Panel Approach." *Journal of Peace Research* 54(6):777–790. doi: 10.1177/0022343317707569.

Yu, J., R. de Jong, and L.-F. Lee. 2008. "Quasi-Maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T Are Large." *Journal of Econometrics* 146:118–134. doi: 10.1016/j.jeconom.2008.08.002.

Zakrewicz, C., B. W. Brorsen, and B. C. Briggeman. 2012. "Comparison of Alternative Sources of Farmland Values." *Agricultural Finance Review* 72(1):68–86. doi: 10.1108/00021461211222132.

Journal of Agricultural and Resource Economics 49(1):S1–S12 ISSN: 1068-5502 (Print); 2327-8285 (Online)

doi: 10.22004/ag.econ.337550

Online Supplement: Spatial and Temporal Spillovers in US Cropland Values

J. Wesley Burnett, Donald J. Lacombe, and Steven Wallander

The Relationship between Farmland Cash Rental Rates and Farm Returns

Figure S1 offers two scatterplot diagrams illustrating the relationship between farm returns and cash rental rates. The returns are measured as net cash farm income and are defined at the county level according to the Census of Agriculture for census years 2012 and 2017 (US Department of Agriculture, 2022). We matched census counties with non-irrigated cash rental rates using US Department of Agriculture (2022) survey data for the same years of observation.

Figure S1(a) offers a diagram of the cash rental rates plotted against binned mean values of net farm income with a line of fit and the standard errors in the gray shaded area.[1] Both cash rents and net farm income were converted to natural logs. The graph demonstrates a positive, linear relationship with a slope coefficient that was equal to 0.37.

To ensure that error variances were nearly equal, a scatter plot of the residuals of cash rental rates against net farm income is offered in Figure S1(b). For the second plot, we estimated a regression of rents on income while controlling for the year of observation and clustering the standard errors at the county level. The scatterplot reveals that the residuals seem to be randomly distributed, which suggests that the assumption that the relationship is linear is reasonable.

(a) Binned Scatterplot (b) Residual Plot

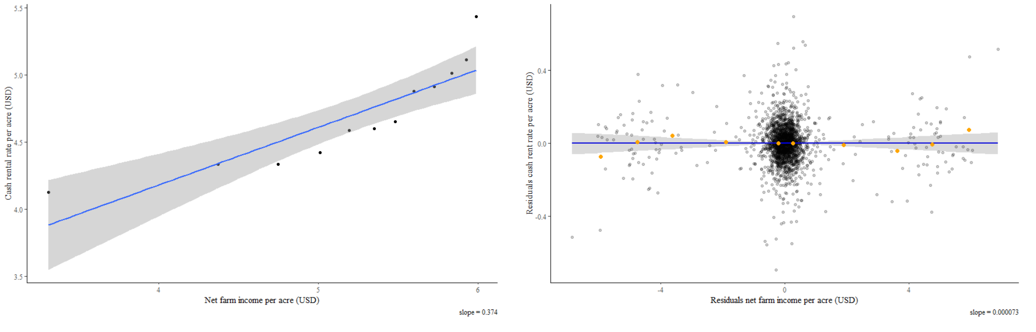

Figure S1. Scatterplots of Cash Rental Rates against Net Farm Income

[1] Figure S1(a) is offered for ease of interpretation of the relationship between rental rates and farm income. Instead of illustrating a scatterplot of thousands of data points (the number of observations), we binned the farm-income data points into 25 separate quantiles. The points in the diagram represent the mean value of the data points within each quantile. It is important to note that the net cash farm income used in this diagram is not the same variable as that used within the regression analyses below.

How to Interpret the Estimates from the Spatial Econometric Models

A major issue in any spatial econometric analysis is the proper calculation of the marginal effects. LeSage and Pace (2009) make note that the partial derivatives in a model with an endogenous peer–effect are not the β's. Instead, the derivative is a matrix of effects estimates. To see this, consider a simple cross–sectional spatial autoregressive model:

$$y = \rho \times Wy + X\beta + \varepsilon. \tag{S1}$$

In (S1), the true partial derivatives are not the β's and this can be seen by writing equation (S1) in reduced form as follows:

$$y = (I_N - \rho W)^{-1}(X\beta) + (I_N - \rho W)^{-1}\varepsilon. \tag{S2}$$

Taking the partial derivative of y with respect to the explanatory variable X in equation (S2) is now equal to $(I_N - \rho \times Wy)^{-1}(\beta_k)$ for all explanatory variables k, which produces a matrix of partial derivatives, one for each point in space. Since $(I_N - \rho \times Wy)^{-1}$ is an $N \times N$ full matrix, a more concise representation of the partial derivative effects is needed. LeSage and Pace (2009) introduce what they call the average direct, indirect, and total effects. The direct effects measure how changes in explanatory variable at location i affect the dependent variable at location i, while the indirect or spillover effects measure how changes in explanatory variables at location j affect the dependent variable at location i, where $i \neq j$.

The effects' estimates in a dynamic spatial panel data model are similar to the effects estimates for a cross–sectional or panel data model with one exception. The difference is that the effects estimates can be decomposed into short–term and long–term effects. Mathematically, we can write the short–term effects estimates in the following manner (Yesilyurt and Elhorst, 2017)

$$\begin{bmatrix} \frac{\partial E(Y)}{\partial x_{1k}} & \cdots & \frac{\partial E(Y)}{\partial x_{Nk}} \end{bmatrix}_t = (I_N - \rho \times W)^{-1}\left[\beta_k I_N\right]. \tag{S3}$$

Equation (S3) is the matrix of partial derivatives of Y with respect to the k^{th} explanatory variable contained within X, which is termed by Elhorst (2014) as the short-run effect. These effects estimates can be decomposed into the direct and indirect or spillover effects. The direct effects are the average of the diagonal elements in equation (S3) and measure how a change in an explanatory variable in the own region affects the dependent variable in the own region, while the average of the row–sums in (S3) are the indirect or spillover effects. The total effects are the sum of the direct and indirect effects.

The long–run effects estimates can be mathematically stated in the following formula (Yesilyurt and Elhorst, 2017)

$$\begin{bmatrix} \frac{\partial E(Y)}{\partial x_{1k}} & \cdots & \frac{\partial E(Y)}{\partial x_{Nk}} \end{bmatrix} = \left[(1-\tau)I_N - (\rho + \eta)W\right]^{-1}\left[\beta_k I_N\right] \tag{S4}$$

Equation (S4) is the matrix of partial derivatives of Y with respect to the k^{th} explanatory variable X of a permanent one unit change in X for each of the N regions through all times T. As in the case of the short–term effects, these effects estimates can be decomposed into the direct and indirect or spillover effects. The direct effects are the average of the diagonal elements in equation (S4) and measure how a change in an explanatory variable in the own region affects the dependent variable in the own region, while the average of the row–sums in (S4) are the indirect or spillover effects. Further details and derivations of these effects estimates can be found in Elhorst (2014) and Debarsy, Ertur, and LeSage (2012).

The dynamic, spatial panel data model used in the current study is based on the methodology proposed by Lee and Yu (2012). The Lee and Yu (2012) method uses quasi-maximum likelihood estimation after eliminating both region-specific and time-specific fixed effects (LeSage, Chih, and Vance, 2019). The spatial fixed effect is eliminated by using an ortho-normal transformation matrix, and the time-specific fixed effects are eliminated by transforming the data to deviations from the time trend (LeSage, Chih, and Vance, 2019). Estimation of the model parameters involves minimizing the negative of the log-likelihood function, subject to stability restrictions on the spatiotemporal parameters.

Spatial Weights Matrices Defined

For this particular study, we chose three pre-specified matrix designs: a contiguity matrix, a distance-based matrix, and a nearest neighbors design. The contiguity matrix is designed such that any two counties that share a common border are assigned a binary unit of one; and, non-neighboring counties are assigned zero. That is, the elements or individual weights (w) within a contiguity, spatial weights matrix, define the spatial relationship between counties i and j $(i \neq j)$ as follows:

$$w_{ij} = \begin{cases} 1, & \text{if } j \in V(i) \\ 0, & \text{if } j \notin V(i), \end{cases} \tag{S5}$$

where $V(i)$ is the boundary of the i^{th} county.

The elements of the distance-based matrix are defined as

$$w_{ij} = \begin{cases} 1, & \text{if } d_{ij} \leq D_m \\ 0, & \text{if } d_{ij} > D_m, \end{cases} \tag{S6}$$

where D_m represents a pre-designated Euclidean distance between centroids of counties i and j, d_{ij}. We defined the pre-designated distance as 100 miles.

The nearest neighbor design selects a pre-defined number of neighboring counties based on the Euclidean distance between centroids. For this study, we specified a design with the six nearest neighboring counties within the radius of an observed county. This can be defined as in equation (S5), where the spatial weight unit is defined as unity if a county is one of the six nearest neighbors to county i – in this context, $V(i)$ is a function that defines the six nearest neighbors.

Evaluation of Multiple Imputation Methods

To address the issue of missing data within the cash rental rates and agricultural land values, we evaluated different multiple imputation methods. All of the multiple imputation methods were estimated using the R package *mice* (multiple imputation by chained equations) (van Buuren and Groothuis-Oudshoorn, 2011).

Before proceeding with imputation methods, it was important to first examine the missing data for any systematic patterns between the missing and observed values. Table S1 provides a summary of the observed and missing values, and an estimate of the percentage of missing values within the dataset. The table only provides estimates for variables that contained missing values. Based on the table, cash rental rates contained the most amount of missing values, followed by the price per acre of farmland based on private transaction records. As indicated in the study, the cash rents are provided by NASS, who often suppresses observations due to potentially revealing personal identifiable information about operators. The farmland price data is missing due to the “thin” market problem described in the introduction. Unfortunately, we were unable to observe a transaction in each year for some of the observed counties. The Conservation Reserve Program county-level acreage data is provided by the Farm Service Agency within the U.S. Department of Agriculture. Unfortunately, this data also contained about 28 percent missingness during this study’s period of observation.

Table S1. Summary Table of Missing Values within the Data

Variable	No. of obs.	No. of missing	Percent missing
Cash rental rate	6,505	8,003	55.2
Cash receipts	14,430	78	0.5
Price per acre	9,140	5,368	37.0
CRP acres	10,398	4,110	28.3

Notes: The second column lists the number of observed values within the data. The third column offers the number of missing values; and the fourth column provides the percentage of missing values. "CRP" denotes the Conservation Reserve Program. The other variables, used within the study, are not provided here as the others do not contain missing values.

To examine any patterns of missingness, Figure S2 demonstrates the observed data relative to the missing data. The light blue color denotes missing observations whereas dark blue denotes observed data. If we are concerned about systematic missingness, which would arguably invalidate the imputation methods (van Buuren, 2018), we would see common patterns looking across the map from left to right or similar patterns of missing data between any two corresponding variables (i.e., looking from up to down the map, or vice versa). The map seems to indicate that the unobserved values are missing at random. Although further investigation (which is the beyond the scope of the current study) would be necessary to determine if the data is missing, not at random (van Buuren, 2018). An example of missing, not a random may include NASS suppressing rental rate observations consistently across the same years and counties. The missing values within Figure S2 do not imply such a case. We leave additional such analyses to future research.

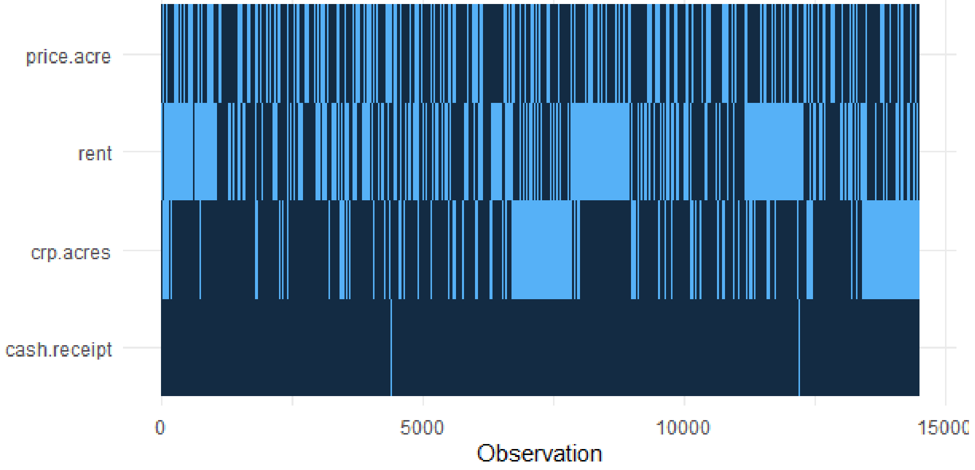

Figure S2. Missing Value Plot

Since the missing data seem missing a random, we proceeded to impute the data using the multiple imputation methods. The methods consisted of predictive mean matching, estimating the mean values, random forest, and a regression-based method. All of the imputation methods were carried out with five separate imputations for each of the variables with significant missingness (cash rents, cash receipts, price per acre, and CRP acres). From the five separate imputations, the method estimates the mean predicted value for a missing entry. Each of the imputation methods were conducted so that we accounted for the fixed effects within the underlying data. In other words, we imputed the missing values using a county classification, so that the missing entry was imputed based on its occurrence within its specified county class. For the sake of space, we do not present all of the results here, but instead offer diagrams for the two imputation methods that provided the best fit to the data. These two methods were the predictive mean matching and random forest.

Predictive mean matching calculates the predicted value of a target variable according a small set of candidate observations from all complete cases that have predicted values closest to the predicted value of the missing entry (van Buuren, 2018). The random forest imputes missing values through classification and regression trees. The random forest method seeks predictors and cut points in the predictors that are used to split the sample into homogeneous subsamples (van Buuren, 2018). The splitting process is repeated on subsample so that a series of splits defines a binary tree, from which a missing entry can be predicted. The random forest method is attractive for imputation as it is robust to outliers, it can deal with multicollinearity and skewness, and it can fit interactions and nonlinear relationships (van Buuren, 2018).

To illustrate how these two methods provided the best fit to the observed data, we provide a set of diagrams showing the observed and imputed distributions of the data. The diagrams for the random forest method are offered in Figure S3 and the diagrams for the predictive mean matching method are offered in Figure S4. As illustrated in Figure S3, the random forest method provided reasonably good fits to the data, but the random forest algorithm had a tendency to fit a symmetric distribution about the mean of the missing variable. This algorithm provided a good fit to the distribution of observed rental data, but unfortunately the shape of imputed distribution differed quite a bit from the observed CRP acres. As can be gleaned in Figure S4, the predictive mean matching method provided nearly identical fits to observed cash rental rates, farmland prices per acre, and CRP acres. The predictive mean matching imputations for cash receipts had a tendency to overestimate the values on the left-hand side of the density, but otherwise, the general shape of the density of the imputed cash receipts is very similar to the observed values.

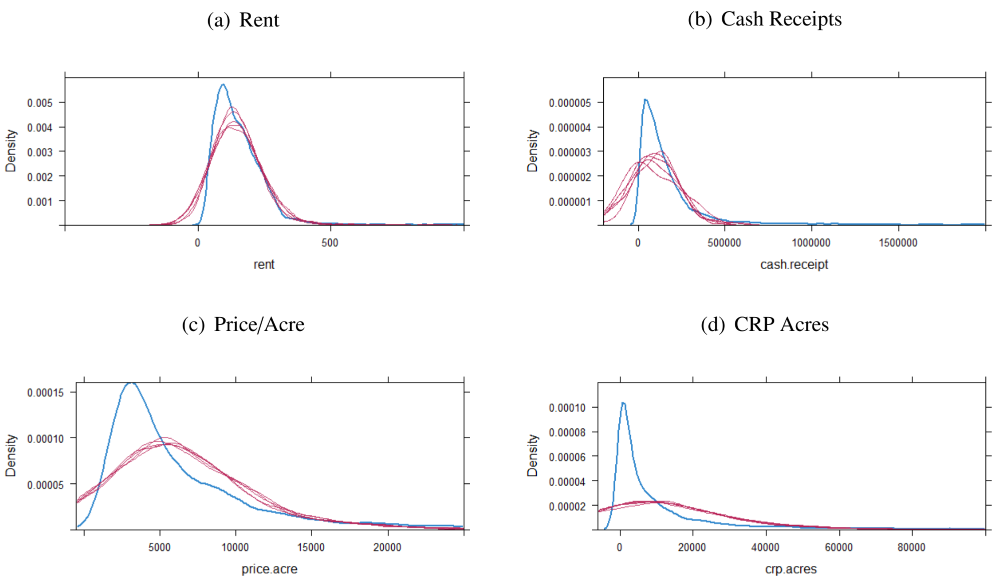

Figure S3. Graphs Showing the Fit of the Random Forest Imputation Method

Notes: The blue lines denote the observed values and the red lines denote the imputed values. There are multiple red lines because the method calculated the missing values with five separate imputations.

Based on the density plots in Figure S3 and S4, we chose to use the imputed values from the predictive mean matching algorithm. The comparisons in distributions for the observed versus the imputed values are offered in Table S2. As demonstrated in the table, there were very small differences in the distributions of the observed and imputed values.

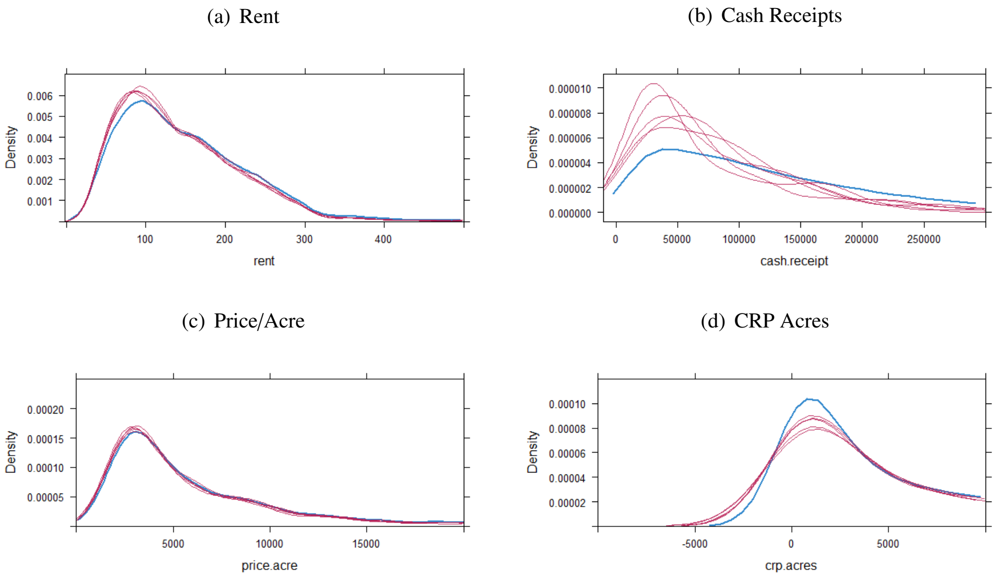

Figure S4. Graphs Showing the Fit of the Predictive Mean Matching Imputation Method

Notes: The blue lines denote the observed values and the red lines denote the imputed values. There are multiple red lines because the method calculated the missing values with five separate imputations.

Table S2. Comparison of Distributions between Non-Imputed and Imputed Variables

Distribution	Rent		Cash Receipts		Price per acre		CRP acres	
Measure	**Observed**	**Imputed**	**Observed**	**Imputed**	**Observed**	**Imputed**	**Observed**	**Imputed**
Minimum	16.0	16.0	2894	2894	0.00	0.00	0.0	0.0
1st quartile	89.0	86.0	53,737	53,453	2,962.08	2,897.15	890.2	865.8
Median	135.0	130.0	110,332	109,919	4,665.50	4,566.80	4,384.7	4,343.3
Mean	168.1	158.0	201,471	200,780	6,575.85	6,444.66	15,022.7	15,348.2
3rd quartile	200.0	193.0	207,240	206,668	8,303.30	8,137.46	14,236.7	14,409.1
Maximum	2610.0	2610.0	6,750,876	6,750,876	73,015.12	73,015.12	276,808.2	276,808.2

Notes: CRP acres denote the amount of acres within the county participating in the Conservation Reserve Program. All distribution estimates are based on the predictive mean matching method with five imputations.

Spatial Fixed Effects Estimates: Distance-Based Spatial Weights Matrix

Table S3. County-Level, Spatial Fixed Effects Estimates

Dependent variable: log(Average agricultural land values)			
Variable	**Spatial Lag** **1**	**Spatial Error** **2**	**Spatial Lag and Error** **3**
Coefficient estimates			
log(Cash rent)	0.1254***	0.1255***	0.1190***
	(0.0109)	(0.0110)	(0.0103)
log(Cash receipts)	0.0323	0.0347	0.0541*
	(0.0318)	(0.0322)	(0.0234)
log(CRP acres)	−0.0083**	−0.0083**	−0.0065*
	(0.0031)	(0.0031)	(0.0028)
log(Population)	0.0404	0.0365	0.0304
	(0.1278)	(0.1313)	(0.1007)
Av. temperature	0.0059	0.0062	0.0010
	(0.0054)	(0.0058)	(0.0027)
Av. precipitation	0.0147	0.0155	0.0122
	(0.0090)	(0.0095)	(0.0064)
Median age	−0.0001	0.0000	−0.0026
	(0.0050)	(0.0051)	(0.0036)
Spatial autocorrelation coefficients			
Spatially lagged dependent	0.0648***		0.3614***
	(0.0125)		(0.0253)
Spatially lagged error		0.0641**	−0.3328***
		(0.0126)	(0.0311)
R^2	0.4620	0.4606	0.4539

Notes: The agricultural land values, rental rate, and farm income data were converted to real values using the GDP deflator. The asterisks terms represent the following: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Estimated standard errors are offered in parentheses. These regression models are based on a distance-based spatial matrix design. “CRP" denotes the Conservation Reserve Program.

Spatial Fixed Effects Estimates: Nearest Neighbors-Based Spatial Weights Matrix

Table S4. County-Level, Spatial Fixed Effects Estimates

Dependent variable: log(Average agricultural land values)			
	Spatial Lag	**Spatial Error**	**Spatial Lag and Error**
Variable	**1**	**2**	**3**
Coefficient estimates			
log(Cash rent)	0.1259***	0.1263***	0.1247***
	(0.0110)	(0.0110)	(0.0109)
log(Cash receipts)	0.0324	0.0335	0.0661*
	(0.0318)	(0.0322)	(0.0258)
log(CRP acres)	−0.0083**	−0.0083**	−0.0071*
	(0.0031)	(0.0031)	(0.0030)
log(Population)	0.0381	0.0329	0.0300
	(0.1279)	(0.1312)	(0.1165)
Av. temperature	0.0060	0.0064	0.0011
	(0.0054)	(0.0058)	(0.0031)
Av. precipitation	0.0060	0.0154	0.0144
	(0.0054)	(0.0095)	(0.0075)
Median age	−0.0001	0.0002	−0.0031
	(0.0050)	(0.0051)	(0.0040)
Spatial autocorrelation coefficients			
Spatially lagged dependent	0.0616***		0.2188**
	(0.0147)		(0.0719)
Spatially lagged error		0.0611***	−0.1718*
		(0.0147)	(0.0874)
R^2	0.4615	0.4606	0.4601

Notes: The agricultural land values, rental rate, and farm income data were converted to real values using the GDP deflator. The asterisks terms represent the following: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Estimated standard errors are offered in parentheses. These regression models are based on a nearest-neighbor spatial matrix design. “CRP" denotes the Conservation Reserve Program.

Direct, Indirect, and Total Impact Estimates: Distance-Based Spatial Weights Matrix

Table S5. Direct, Indirect, and Total Impact Estimates Based on Dynamic, Spatial Panel Data Models

Dependent variable: log(Agricultural land values)			
	Direct	**Indirect**	**Total**
Short-run impacts			
log(Cash rents)	0.1175***	0.0067***	0.1242***
	(0.0121)	(0.0020)	(0.0129)
log(Cash receipts)	0.0233	0.0013	0.0247
	(0.0364)	(0.0022)	(0.0385)
log(CRP acres)	−0.0064*	−0.0004	−0.0067*
	(0.0032)	(0.0002)	(0.0034)
log(Population)	0.1184	0.0066	0.1250
	(0.1366)	(0.0085)	(0.1445)
Average temp	0.0056	0.0003	0.0059
	(0.0053)	(0.0003)	(0.0056)
Long-run impacts			
log(Cash rents)	0.1228***	0.0113***	0.1341***
	(0.0127)	(0.0034)	(0.0141)
log(Cash receipts)	0.0244	0.0022	0.0266
	(0.0380)	(0.0038)	(0.0416)
log(CRP acres)	−0.0067*	−0.0006	−0.0073*
	(0.0034)	(0.0004)	(0.0037)
log(Population)	0.1237	0.0111	0.1348
	(0.1426)	(0.0038)	(0.1557)
Average temp	0.0059	0.0005	0.0064
	(0.0055)	(0.0006)	(0.0060)

Notes: All impact estimates are based on the dynamic, spatial panel data estimator with a distance-based spatial weights matrix. The agricultural land values, rental rates, and farm income data were converted to real values using the implicit GDP deflator. The asterisks terms represent the following: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Estimated standard errors are offered in parentheses. Other variables not included average precipitation and median age within the county – both were statistically insignificant. “CRP" denotes the Conservation Reserve Program.

Direct, Indirect, and Total Impact Estimates: Nearest Neighbor-based Spatial Weights Matrix

Table S6. Direct, Indirect, and Total Impact Estimates Based on Dynamic, Spatial Panel Data Models

Dependent variable: log(Agricultural land values)			
	Direct	**Indirect**	**Total**
Short-run impacts			
log(Cash rents)	0.1178^{***}	0.0056^{*}	0.1233^{***}
	(0.0122)	(0.0023)	(0.0129)
log(Cash receipts)	0.0219	0.0010	0.0229
	(0.0364)	(0.0020)	(0.0381)
log(CRP acres)	-0.0064^{*}	−0.0003	-0.0067^{*}
	(0.0032)	(0.0002)	(0.0034)
log(Population)	0.1155	0.0053	0.1208
	(0.1367)	(0.0075)	(0.1433)
Average temp	0.0058	0.0003	0.0061
	(0.0053)	(0.0003)	(0.0055)
Long-run impacts			
log(Cash rents)	0.1231^{***}	0.0131^{**}	0.1362^{***}
	(0.0127)	(0.0042)	(0.0145)
log(Cash receipts)	0.0229	0.0024	0.0253
	(0.0380)	(0.0069)	(0.0422)
log(CRP acres)	-0.0067^{*}	−0.0007	-0.0074^{*}
	(0.0034)	(0.0004)	(0.0037)
log(Population)	0.1207	0.0125	0.1332
	(0.1428)	(0.0165)	(0.1580)
Average temp	0.0061	0.0007	0.0067
	(0.0055)	(0.0007)	(0.0061)

Notes: All impact estimates are based on the dynamic, spatial panel data estimator with a nearest-neighbor spatial weights matrix. The agricultural land values, rental rates, and farm income data were converted to real values using the GDP deflator. The dot and asterisks terms represent the following: $^{\cdot}$ $p < 0.1$; * $p < 0.01$; ** $p < 0.001$; *** $p < 0$. Estimated standard errors are offered in parentheses. Other variables not included are median age within the county and average precipitation – neither variable was statistically significant. “CRP" denotes the Conservation Reserve Program. Impact estimates for the other two spatial weights matrix specifications are offered in the online appendix.

Calculations of the Forecasting Error Metrics

In order to carry out the forecasts, we defined the first ten years (2008-2017) of observations as the training set and the last three years (2018-2020) as the test set. In other words, we estimated the dynamic, spatial panel data model using the training data and then predicted the per-acre, farmland sales prices for the last three years of data. One could interpret the in-sample as the training data set and the out-of-sample as the test data set. To calculate the forecasted per-acre prices, we multiplied the coefficient estimates from the training-data regressions by the explanatory variables in the test data set.

After estimating the predicted values, we calculated five separate error forecasting metrics. The metrics included: the mean absolute error (MAE), root mean squared error (RMSE), normalized root mean squared error (NRMSE), weighted average percentage error (WAPE), and the mean absolute percentage error (MAPE). Each of these metrics evaluates the forecasting errors (the out-of-sample equivalent to the in-sample residuals) based on different assumptions of the data generating process (Allen and Fildes, 2001). As a robustness check, we provide the estimates for all five of the metrics.

The mean absolute error was calculated as

$$\text{(S7)} \qquad \text{MAE} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}|X_{it} - \hat{X}_{it}|}{N \times T},$$

where T and N denote the total number of time and regional (counties) observations. Moreover, the terms X_{it} and $\hat{X}_{it}$ are defined as the actual (observed) values and the forecasted values, respectively.

The root mean squared error was calculated as

$$\text{(S8)} \qquad \text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \hat{X}_{it})^2}{N \times T}}.$$

We defined the normalized, root mean squared error as

$$\text{(S9)} \qquad \text{NRMSE} = \frac{RMSE}{\bar{X_{it}}},$$

where $\bar{X_{it}}$ denotes the mean of the observed per-acre sales price of agricultural land.

The weighted average percentage error was defined as

$$\text{(S10)} \qquad \text{WAPE} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}|X_{it} - \hat{X}_{it}|}{\sum_{i=1}^{N}\sum_{t=1}^{T}|X_{it}|}.$$

Finally, we defined the mean absolute percentage error as

$$\text{(S11)} \qquad \text{MAPE} = \frac{1}{N \times T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\frac{X_{it} - \hat{X}_{it}}{X_{it}}\right|.$$

References

Allen, P. G., and R. Fildes. 2001. "Econometric Forecasting." In J. Armstrong, ed., *Principles of Forecasting: A Handbook for Researchers and Practitioners, Technological Forecasting and Social Change*, Vol. 69(3). New York, NY: Springer, 313–316. doi: 10.1016/S0040-1625(02)00180-4.

Debarsy, N., C. Ertur, and J. P. LeSage. 2012. "Interpreting Dynamic Space–Time Panel Data Models." *Statistical Methodology* 9(1-2):158–171.

Elhorst, J. P. 2014. "Dynamic Spatial Panels: Models, Methods and Inferences." In *Spatial Econometrics,* Springer, 95–119. doi: 10.1007/s10109-011-0158-4.

Lee, L.-F., and J. Yu. 2012. "QML Estimation of Spatial Dynamic Panel Data Models with Time Varying Spatial Weight Matrices." *Spatial Economic Analysis* 7(1):31–74.

LeSage, J., and R. K. Pace. 2009. *Introduction to Spatial Econometrics*. Chapman and Hall/CRC.

LeSage, J. P., Y.-Y. Chih, and C. Vance. 2019. "Markov Chain Monte Carlo Estimation of Spatial Dynamic Panel Models for Large Samples." *Computational Statistics & Data Analysis* 138: 107–125.

US Department of Agriculture. 2022. "Quick Stats." Washington, DC: USDA National Agricultural Statistics Service. Available online at https://quickstats.nass.usda.gov/ [Accessed June 2022].

van Buuren, S. 2018. *Flexible Imputation of Missing Data*, 2nd ed. New York, NY: CRC.

van Buuren, S., and K. Groothuis-Oudshoorn. 2011. "mice: Multivariate IMputation by Chained Equations in R." *Journal of Statistical Software* 45(3):1–67. doi: 10.18637/jss.v045.i03.

Yesilyurt, M. E., and J. P. Elhorst. 2017. "Impacts of Neighboring Countries on Military Expenditures: A Dynamic Spatial Panel Approach." *Journal of Peace Research* 54(6):777–790. doi: 10.1177/0022343317707569.