

Hot Spots and Spatial Autocorrelation in Certified Organic Operations in the United States

I. Julia Marasteanu, and Edward C. Jaenicke

This study identifies clusters of certified organic operations in the United States and determines the form of spatial autocorrelation present in the operations' distribution. We identify large hot spots of organic operations along the West Coast and in the Midwest and Northeast with some variation based on how we define an organic operation. Further analyses suggest that organic operations do not necessarily follow the same geographic patterns as nonorganic agricultural and general business establishments. Spatial autoregressive models confirm the presence of significant spatial dependence in the distribution of certified organic operations for a number of different definitions of an organic operation.

Key Words: agricultural economics, cluster, hot spot, local Moran's I, organic agriculture, spatial econometrics

Due to a wide variety of natural and economic factors, agriculture is a highly spatial industry. Existing analyses often reveal the presence of both spatial autocorrelation, which reflects the concept that agricultural processes are correlated over distance or across geographic boundaries, and hot spots, which are clustered areas with positively correlated high-attribute values (discussed further later). Roe, Irwin, and Sharp (2002) found that spatial lags were present in data relating to the hog production industry. Schmit and Hall (2013) addressed the presence and impact of food industry clusters while Stewart et al. (2009) analyzed clusters specifically in Tennessee's agriculture industry. Grogan and Goodhue (2012) used spatial analysis and found evidence that citrus producers chose pest control methods similar to those of nearby producers. Identification of hot spots or industry clusters is important because research often shows that they can be advantageous for economic development (Morrison Paul and Seigel 1999, Feser 1998, Chevassus-Lozza and Galliano 2003, Cainelli 2008, Glaeser et al. 1992, Greenstone, Hornbeck, and Moretti 2010, Barkley and Henry 1997, Duranton and Puga 2004, Gibbs

I. Julia Marasteanu is an economist with the U.S. Food and Drug Administration Edward C. Jaenicke is a professor in the Department of Agricultural Economics, Sociology, and Education at Penn State University. *Correspondence:* Julia Marasteanu - U.S. Food and Drug Administration - 5100 Campus Drive - College Park, MD 20740 - Phone 240.402.6159 - Email juliamarasteanu@gmail.com, Ioana.Marasteanu@fda.hhs.gov.

The views expressed here are those of the authors and not necessarily those of the U.S. Food and Drug Administration.

Agricultural and Resource Economics Review 45/3 (December 2016) 485–521

© The Author(s) 2016. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

and Bernat 1997, Gabe 2004, 2008, Graham and Kim 2008, Rocha and Sternberg 2005, Feser, Renski, and Goldstein 2008).

No theoretical reason has been found to suggest that hot spots and spatial correlation in organic agriculture, which we define as including organic production and handling, should be similar to agriculture in general. As a special case of agriculture, organic operations display different characteristics from those of conventional agricultural operations, including more restricted production methods (National Organic Program 2015), higher input costs (Economic Research Service 2012), need for more specialized labor (Klonsky and Tourte 1998), and more-frequent use of their own resources (Argiles and Brown 2010, Schmidtner et al. 2012). The demand side of the organic food market is different as well. Consumers of organic tend to have stronger concerns about healthiness, the environment, food safety, animal welfare, and local economies (Hughner et al. 2007). These factors imply that the needs of organic operations are different and that such operations will not necessarily gravitate only to areas where agricultural establishments are prevalent.

The organic food industry is growing at a much faster rate than the conventional food industry; retail sales of organic food increased from \$11 billion in 2004 to \$27 billion in 2012 (Osteen, Gottlieb, and Vasavada 2012, Onken, Bernard, and Pesek 2011), and the amount of land devoted to organic food crops doubled between 1997 and 2005 (Dimitri and Oberholtzer 2009). These trends have led to concern that some sectors of the organic food industry (particularly corn and soybeans, which are important inputs in production of organic milk and dairy products) may not be growing at a sufficiently fast pace to keep up with the increasing demand (Dimitri and Oberholtzer 2009).

Because of the differences between organic and conventional agriculture and the growth of specific sectors of the organic industry, it is important to analyze and better understand where organic operations are thriving and where they are not. More specifically, understanding spatial relationships in the concentration of organic operations and how organic and nonorganic spatial relationships potentially differ may prove important as policymakers and industry participants plan for continued growth. This type of spatial analysis and identification of organic hot spots or cold spots is also an important first step in further research on the organic sector. Once the hot and cold spots are identified, that information can be used to answer follow-up questions relating to their formation and the economic impact of organic hot spots. This type of spatial analysis can therefore be valuable to governmental and private organizations that focus on development of the organic sector and on regional economic development.

Documenting spatially defined hot spots and spatial correlation in organic agriculture generally has not been done for U.S. organic agriculture, though Schmidtner et al. (2012) did for German organic agriculture. A *New York Times* article (Fairfield 2009) mentioned clustering in the U.S. organic sector but did not use spatial methods to identify the clusters, while Eades and

Brown (2006) used spatial methods to identify clusters but did not identify cold spots. Neither of these sources analyzed the type of spatial autocorrelation. Furthermore, to our knowledge, prior research on organic clustering has not employed a bivariate analysis to compare types of organic operations or organic clusters to other types of clusters and has not examined whether (and if so, how) the size and location of organic clusters has changed over time.

This study systematically investigates clustering as it pertains to the U.S. organic agricultural sector. We first identify clusters of certified organic operations, analyze how they vary across primary scope (crops, handling, or livestock), and compare the locations to those of clusters of agricultural establishments and clusters of all establishments, including those not in agriculture. To assess if and how the clusters evolved over time, we compare organic clusters in 2009 and 2013. We then determine the form of spatial autocorrelation present in the distribution of organic operations by estimating spatial econometric models. In essence, we address the following questions: Where are county-level organic hot spots and cold spots? How might the hot spot and cold spot patterns differ if organic operations are defined more narrowly by primary scope or operation type? How have organic hot spots and cold spots evolved in recent years? How do the locations of organic hot spots and cold spots differ from those of conventional agriculture or establishments in general? What form does spatial autocorrelation take in the county-level distribution of certified organic operations?

Participants and observers may have speculated about hot spots, cold spots, or spatial autocorrelation, but our results provide one of the first efforts to document them using well-defined spatial statistical methods applied to the directory of certified organic operations collected by the U.S. Department of Agriculture's (USDA's) National Organic Program. Our cluster analysis of hot spots and cold spots is set apart from prior studies by its (i) use of a National Organic Program data set, (ii) analysis of different types of organic operations, including organic handlers, (iii) side-by-side comparisons and bivariate analyses of how different types of organic, agricultural, and general establishment clusters are related, and (iv) analysis of the evolution of organic clusters over time.

Our research identifies statistically significant hot spots for certified organic operations in parts of the West Coast, the Midwest, and the Northeast. The locations of these organic hot spots often do not coincide with hot spots of conventional agricultural establishments or general establishments. For example, some areas that contain a relatively small number of general and agricultural establishments are surrounded by areas with relatively high numbers of organic operations, and vice versa. We find that the overall clustering of organic operations remained stable between 2009¹ and 2013.

¹ To our knowledge, National Organic Program lists of organic operations that were certified and exempt prior to 2009 are not publicly available.

Our results further suggest the presence of spatial autocorrelation and cross-county spillovers and that the type and level of spatial correlation and the geographic distribution of hot spots vary with the definition of organic operation used. The results generally are robust across different spatial econometric models.

Background on Cluster Identification and the Organic Food Sector

Much of the research on clusters (e.g., Porter 2003, Goetz, Shields, and Wang 2008) has focused on their identification and the definition of the term “cluster.” Some studies defined clusters based solely on geographic proximity of the industries or firms of interest (e.g., Lian et al. 2009, Banasick, Lin, and Hanham 2009, Barkley and Henry 1997, Eades and Brown 2006), some solely on economic interaction (e.g., Feser and Bergman 2000), and others on both geographic proximity and economic interaction (e.g., Greenstone, Hornbeck, and Moretti 2010, Rocha and Sternberg 2005).

The studies that used geographic proximity employed various mathematical methods to test for and identify spatial autocorrelation, which measures the degree to which observations depend on the characteristics of neighbors (LeSage 1998, Anselin 1999). The Moran’s I and local Moran’s I (LMI) test statistics, which can both be used to test the null hypothesis of no spatial autocorrelation, are commonly used (e.g., Lian et al. 2009, Banasick, Lin, and Hanham 2009, Moons, Brijs, and Wets 2008, Zhang and Lin 2008, Eades and Brown 2006, Richards, Hamilton, and Patterson 2010, Hatzenbuehler, Gillespie, and O’Neil 2012, Schmidtner et al. 2012). In a study of organic agriculture, Schmidtner et al. (2012), for example, used the LMI to identify hot spots of organic farming in integrated counties in Germany.

In econometric models, spatial autocorrelation can be present in the dependent variables, the independent variables, and the residuals. Spatial autocorrelation in the dependent variable (spatial lag) implies that an observation at one location depends on observations at other locations and, thus, that variation in the dependent variable is due in part to differences at neighboring locations. In that case, the covariance of random variables at two locations is not zero. Spatial autocorrelation in the residuals of a regression (spatial error or heterogeneity) implies that the relationships vary over space, which further suggests nonconstant error variances and heteroskedasticity. Spatial autoregressive models are frequently used (e.g., Schmidtner et al. 2012) to account for the possibility of spatial lag and spatial error in regressions (LeSage 1998, Anselin 1999).

Other methods sometimes used to define geographic clusters include the location quotient, which is a measure of industry concentration (e.g., Greenstone, Hornbeck, and Moretti 2010, Feser et al. 2008, Eades and Brown 2006, Rocha and Sternberg 2005); locational Gini coefficients, which are used to analyze the geographic distribution of employment for a specific industry in a region (e.g., Barkley and Henry 1997); and the Hirschman-Herfindahl

Index (HHI), which is a general measure of concentration (Eades and Brown 2006, Lopez et al. 2002). Studies that defined clusters using economic interaction have typically used data on input-output flows (e.g., Rocha and Sternberg 2005, Stewart et al. 2009).

The definition chosen for a cluster is important. Jaenicke et al. (2009), for example, found that the impact of clustering on the output of organic handling firms varied as the number of firms necessary to constitute a cluster changed. Rocha and Sternberg (2005) found that the effects of clusters defined by geographic proximity were different from the effects of clusters defined by economic interaction.

Some recent research on clustering in the food and agriculture industry has specifically looked at the organic food sector. Eades and Brown (2006), for example, found evidence of clustering in organic agriculture, and Schmidtner et al. (2012) found that agglomeration influenced the spatial distribution of organic farms in Germany. Naik and Nagadevara (2010), in a study of organic farming in Karnatka, India, identified economic benefits from clustering, and Jaenicke et al. (2009), in a study of the United States, found that clustering positively affected the output (in sales per employee) of organic handling firms. Using the location quotient to measure concentration, Taus, Ogneva-Himmelberger, and Rogan (2013) found that spatial dependence was a factor in the spatial distribution of farms in the United States being converted to organic agriculture. Hooker and Shanahan (2012) used the Gini coefficient to measure concentration in the U.S. organic sector and a multiplicative model to analyze the impact that market access and input variables had on dispersion of organic producers.

Methodology: Identification of Hot Spots

To identify statistically significant hot spots (defined as counties with positively correlated, high values for attributes), cold spots (defined as counties with positively correlated, low values for attributes), and outliers (counties with negatively correlated attributes), we use the LMI test statistic. Anselin (1995) provides a clear explanation of the LMI and how it is calculated, and the ArcGIS website (www.arcgis.com) explains the statistic in the context of the software. As previously noted, the LMI is used to test the null hypothesis of no spatial autocorrelation, and it allows one to divide the area of interest into smaller sections to test for the presence of local spatial autocorrelation. The LMI is calculated as

$$I_i = (x_i - X_{avg}) \sum_{(j \neq i)} w_{ij} (x_j - X_{avg})$$

where x_i represents the attribute level for section i , X_{avg} is the mean attribute level for the entire area, and w_{ij} is the weighting value between sections i and j . In our case, the sections are counties, the entire area is the United States,

and the attribute level for county i is the number of organic operations (and, for comparison, the number of total agricultural farm establishments and total general establishments).

Spatial weighting matrices generally are based on contiguity or true distance. Two commonly used distance-based matrices are row-standardized distance bands and row-standardized inverse distances. A distance-band matrix assigns a weight of 1 for two counties that are within a specified threshold distance of each other and 0 otherwise (LeSage 1998, Anselin 1999). In a row-standardized matrix, the elements in each row are standardized to sum to 1. The distance threshold typically is defined as the minimum distance required for every section (county, in our case) to have at least one neighbor (GeoDa Center 2013). An inverse-distance matrix uses the inverse of the distance between two counties as their weight and can be constrained by a distance band (LeSage 1998, Anselin, 1999). To provide additional insight and check the robustness of the results, we compare the sets of clusters generated by using a contiguity weighting matrix and a distance-band weighting matrix.²

To determine the significance of the LMI, we use a permutation method implemented in GeoDa software. The observation, or attribute level, in the county analyzed is held constant and all of the other observations are assigned a value from a vector of random numbers to relocate them in space, producing a random spatial distribution. The LMI for the county of interest is then calculated using the random spatial distribution, and the process is repeated multiple times using different random number seeds. The p-value is based on the probability that the actual LMI for the county is equal to the values calculated during the permutations (GeoDa Center 2013). Unlike z-score tests for significance (e.g., LeSage 1998, Anselin 1999), this method does not compute the significance of the LMI analytically, but it has the advantage of being able to test for the robustness of our results by comparing the results generated by different seeds and varying the number of permutations.³

To better compare the distribution of organic operations to the distributions of agricultural farm establishments and general establishments, we also use the bivariate LMI test statistic (Anselin, Syabri, and Kho 2006). It is similar to the univariate LMI but compares the level of an attribute in an area to the level of a different attribute in neighboring areas (e.g., comparing the number of organic operations in one area to the number of agricultural farm establishments in neighboring areas):

² Cluster maps generated by another program using an inverse-distance weighting matrix yielded similar results; the hot spots were present in the same general areas and there were fewer cold spots and outliers.

³ Results when we varied the seeds and number of permutations were similar.

$$I_i = (x_i - X_{avg}) \sum_{(j \neq i)} w_{ij} (y_j - Y_{avg})$$

where x_i is the level of attribute x for section i , X_{avg} is the mean level of attribute x for the entire area, y_i is the level of attribute y for section i , Y_{avg} is the mean level of attribute y for the entire area, and w_{ij} is the weighting value between sections i and j .

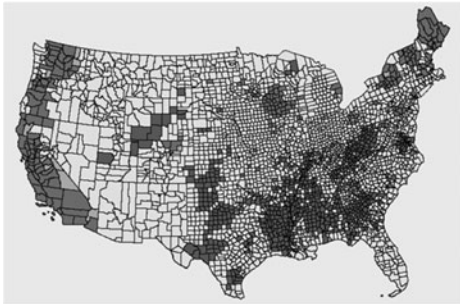
Under the bivariate LMI, interpretation of the hot spots, cold spots, and outliers changes slightly. A hot spot in this case indicates that a high level of an attribute in an area is positively correlated with high levels of another attribute in neighboring areas. Interpretation of the cold spots and outliers is similar. The procedure for determining the significance of the test statistic remains the same.

Our data on certified organic operations come from USDA's National Organic Program and are publicly available online. They consist of a list of all certified organic operations, along with information such as operation name, certifying agent, primary scope (i.e., handling, crops, and livestock), address, phone number, and products produced. Approximately 60 percent of the operations have crops as their primary scope while 28.5 percent have handling, 11.4 percent have livestock, and less than 1 percent have wild crops as their primary scopes. The data on agricultural operations come from the 2007 U.S. Census of Agriculture and includes information on agricultural farms. To facilitate the comparison between organic hot spots and general agricultural hot spots, we also create a variable for organic production operations (crops and livestock) only.

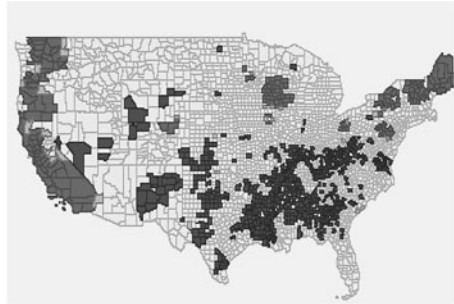
Hot Spot and Cold Spot Results

Figure 1 shows the clusters identified in our analysis for certified organic operations with separate maps of clusters for all organic operations and for organic production (crops and livestock), crops, handling, and livestock calculated using the univariate LMI statistic and a queen contiguity matrix (which considers both shared boundaries and shared points as contiguous).⁴ We find that the clusters for all organic operations and organic production are very similar. The largest hot spot for both runs along nearly the entire West Coast. Another large hot spot is centered in part of Wisconsin and southeastern Minnesota, and others appear in Maine, New Hampshire, Vermont, Massachusetts, upstate New York, and southeastern Pennsylvania. Some relatively small hot spots are found in the Midwest and West. Cold spots, on the other hand, occur mainly in Southern states and cover many areas from Texas to Virginia. There are also relatively small cold spots in the

⁴ Maps generated using a distance-band matrix showed hot spots and cold spots that were larger but were located in the same general areas and a greater number of outliers.



1.1 All Organic Operations: Crops, Livestock, Handling



1.2 Organic Production: Crops and Livestock Only



1.3 Organic Crops



1.4 Organic Handling



1.5 Organic Livestock

Gray = Not Significant
 Red = Hot Spot
 Blue = Cold Spot
 Purple = Low-high
 Pink = High-low

Figure 1. Organic Hot Spots, Cold Spots, and Outliers Based on Operation Counts

West, Midwest, Alaska, and Hawaii. The outliers are scattered throughout the country.

Hot spots for organic crops and handling are similarly situated, mostly along the West Coast and in the upper Midwest and parts of the Northeast and New England. Hot spots for organic livestock, on the other hand, are more isolated and scattered; the only concentrations are in northern New England and New York. We find numerous cold spots for organic crops throughout the South and in parts of the West, Appalachia, and the mid-Atlantic states; a

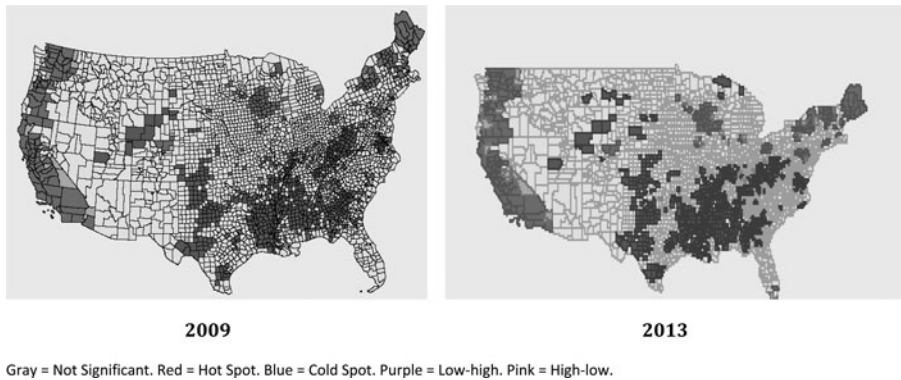


Figure 2. Organic Operation Hot Spots: 2009 and 2013

smaller number of cold spots for organic handlers; and even fewer cold spots for livestock producers.

Figure 2 compares hot and cold spots for all organic operations in 2009 and 2013. We find that the distributions are generally the same, indicating that the organic industry remained relatively stable during those years. One notable difference is apparent growth of organic operations in southern Florida. Among all U.S. counties, 26 were hot spots in 2013 but not 2009 and 12 were hot spots in 2009 but not 2013. Similarly, 156 counties were cold spots in 2013 but not in 2009 and 145 were hot spots in 2009 but not 2013. It is important to note that some of this variation may be related to our use of a permutation method to identify hot spots. Because of a lack of data on organic operations prior to 2009, it is difficult to draw more-general conclusions about the distribution of organic operations over time.

We next identified hot and cold spots for all agricultural establishments using U.S. Census of Agriculture data (2007) and the univariate LMI and then used a bivariate LMI statistic that uses both organic production and agricultural farms data to generate a different form of hot spots and cold spots. Figure 3.1 shows a map of clusters of agricultural production. When compared to Figure 1.2, Figure 3.1 shows that organic production hot spots and general agriculture hot spots match some but not most of the time, so that organic production does not necessarily follow the same pattern as conventional production. Both maps show hot spots along the West Coast and throughout the Midwest. However, many of the organic hot spots in the northeastern states are not present for agricultural farming. Some agricultural production hot spots now appear in Florida, Alabama, Louisiana, Texas, Oklahoma, and Arkansas.⁵

⁵ A potential avenue for future research would be to determine the reasons for these discrepancies, though the ones in the Southeast likely reflect production of crops such as

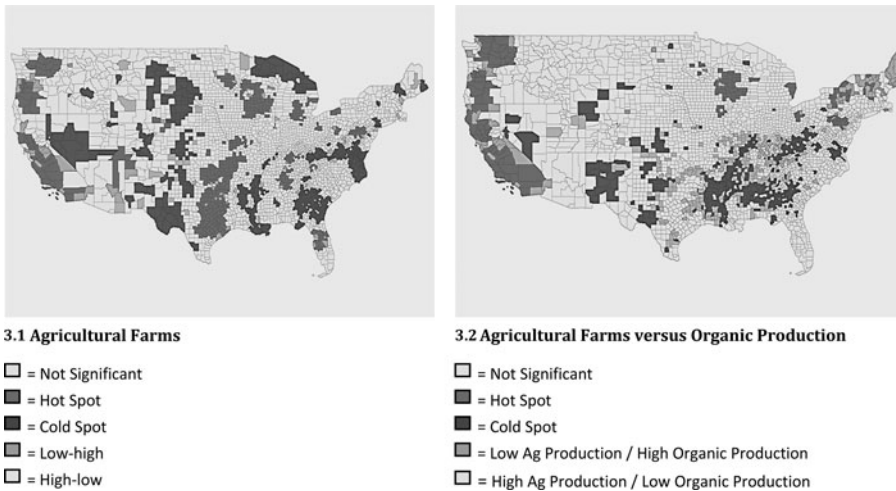


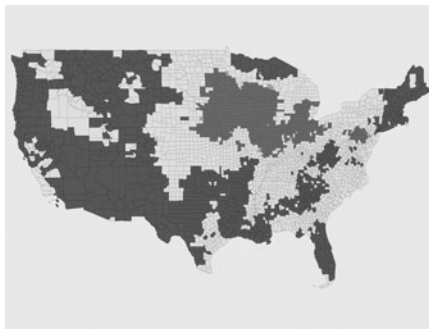
Figure 3. Hot Spots, Cold Spots, and Outliers of Agricultural Farms

Figure 3.2 provides a more formal comparison of organic production and general agricultural production. The definitions of hot spots, cold spots, and outliers are somewhat different because we use the bivariate LMI. We are particularly interested in areas with high-low spots—counties that have a relatively large number of general agricultural establishments surrounded by counties that have a relatively small number of organic production operations—and in areas with low-high spots—counties that have a relatively small number of agricultural production establishments surrounded by counties that have a relatively large number of organic production operations. Many of the high-low areas are in the South and Southeast with others scattered around the rest of the country. The low-high areas are in the Northeast and in the eastern portions of California, Oregon, and Washington. These results show that organic operations do not necessarily concentrate in primarily agricultural areas. In future research, it might be interesting to analyze the factors associated with these discrepancies.

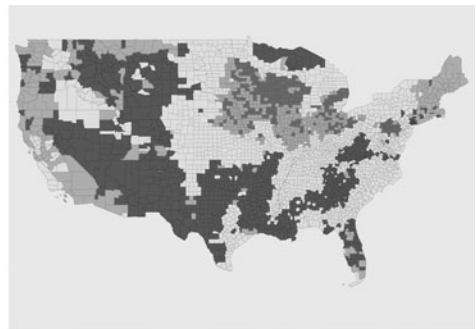
Another aspect of growth in the organic sector is the concern that certain parts of this sector may not be growing at a sufficiently fast rate. To highlight this, Figure 4 looks at one specific type of agricultural crop, corn.⁶ Figure 4.1 shows a map generated by the univariate LMI for operations that sell corn

cotton, which are prevalent in the area (Environmental Protection Agency 2014) and are not typically produced according to organic standards. For example, according to the National Agricultural Statistics Service (Economic Research Service 2008) only 0.14 percent of cotton produced in the United States in 2008 was certified as organic.

⁶ Similar analyses for soybean crops yield very similar results.

**4.1 Corn Production**

Gray = Not Significant
 Red = Hot Spot
 Blue = Cold Spot
 Purple = Low-high
 Pink = High-low

**4.2 Corn Production versus Organic Crops**

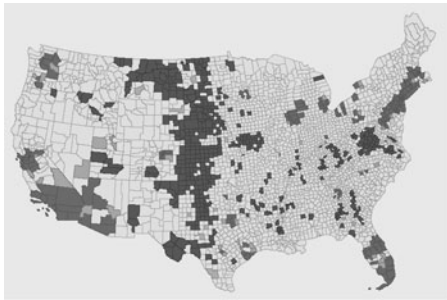
Gray = Not Significant
 Red = Hot Spot
 Blue = Cold Spot
 Purple = Low Corn Production / High Organic Crops
 Pink = High Corn Production / Low Organic Crops

Figure 4. Hot Spots, Cold Spots, and Outliers of Corn Production

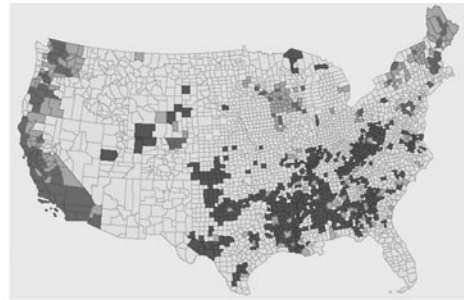
crops. It essentially shows that corn production was concentrated in the Midwest and almost nowhere else. More interesting, however, is [Figure 4.2](#), which shows a bivariate analysis of corn to organic crops. Particularly relevant is the large area of low-high spots in the Midwest—counties that have a relatively low number of corn production operations but are surrounded by counties with high numbers of organic crop production operations. These low-high spots are consistent with concerns about insufficient expansion of organic corn crops discussed in Dimitri and Oberholtzer (2009).

[Figure 5](#) presents the results of a similar comparison of organic establishments versus establishments in general. We use business establishments in general as a proxy for density of economic activity and proximity to urban areas to determine whether organic agriculture thrives only in areas with high levels of economic activity. Comparing [Figures 5.1](#) and [1.1](#), we find that hot spots of business establishments generally are less common than hot spots of organic operations. The largest concentrations of businesses in general are in northern and southern California, the Southwest, Florida, the Northeast corridor, and a multi-state area surrounding Chicago. The bivariate LMI analysis ([Figure 5.2](#)) identifies large areas of low-high spots along the West Coast, in parts of the Midwest, and in the Northeast. Interestingly, these same areas are organic hot spots (see [Figure 1.1](#)).

This result suggests that organic operations do not necessarily gravitate to areas that have a large number of establishments in general; on the contrary, we find that organic operations tend to concentrate in areas that have

**5.1 General Establishments**

Gray = Not Significant
 Red = Hot Spot
 Blue = Cold Spot
 Purple = Low-high
 Pink = High-low

**5.2 General Establishments versus Organic Establishments**

Gray = Not Significant
 Red = Hot Spot
 Blue = Cold Spot
 Purple = Low General Establishments / High Organic Production
 Pink = High General Establishments / Low Organic Production

Figure 5. Hot Spots, Cold Spots, and Outliers of General Establishments

relatively few establishments and would not be characterized by sprawling development (Brown, Goetz, and Fleming 2012). Some of these areas (parts of the Midwest and West Coast) coincide with hot spots of agriculture generally and thus may indicate the presence of large agricultural industries. Figure 5.2 also shows some isolated high-low spots scattered across the country, especially in the South.

Figure 6 shows the results of a bivariate LMI analysis for different types of organic operations. The first panel (6.1) shows the results of a comparison of handling versus production (crops and livestock) and identifies a number of areas in California and the Midwest in which there is a high level of organic production but relatively few handling operations. This pattern could reflect either a reliance on direct market sales in those areas or that organic production in those areas is transported to other areas for primary handling.⁷ The second panel (6.2) compares production of organic crops with handling and suggests that their spatial distributions are similar with few high-low and low-high spots. The last two panels compare organic livestock production to organic crops (6.3) and organic handling (6.4). Many areas scattered throughout the country have a high level of organic crops or handling and few organic livestock operations. The only large area that has a high level of organic livestock and a low level of crops or handling is in the upper Northeast.

⁷ Factors associated with these differences would be interesting to analyze further in future research.

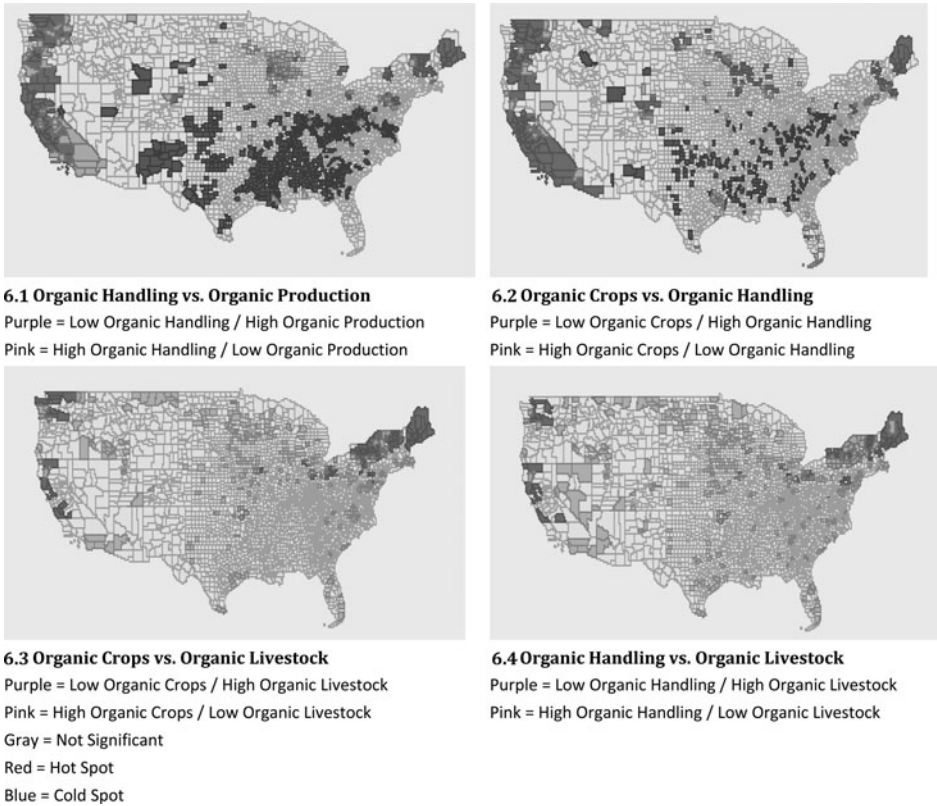


Figure 6. Comparison of Types of Hot Spots based on the Bivariate Local Moran's I

These results provide a starting point for further exploration of factors associated with the presence and formation of hot spots of organic operations. Several studies of the organic sector have addressed factors that promote or inhibit its growth. Farmer et al. (2013) and Farmer (2014), for example, cited challenges and costs associated with sustaining organic growing methods, the cost of USDA certification and producers' cynicism about it, concerns about a lack of benefit from certification (especially among operators of small farms), and producers' and consumers' lack of access to information as factors that inhibited the growth of organic agriculture in West Virginia. They also found that farmers who chose to pursue organic certification tended to be relatively young, had a higher level of education, split work evenly among men and women, operated relatively small farms, and had better access to the internet. On a more general scale, Marasteanu and Jaenicke (*forthcoming*) discuss several factors potentially leading to formation of hot spots of organic agricultural in the United States and find

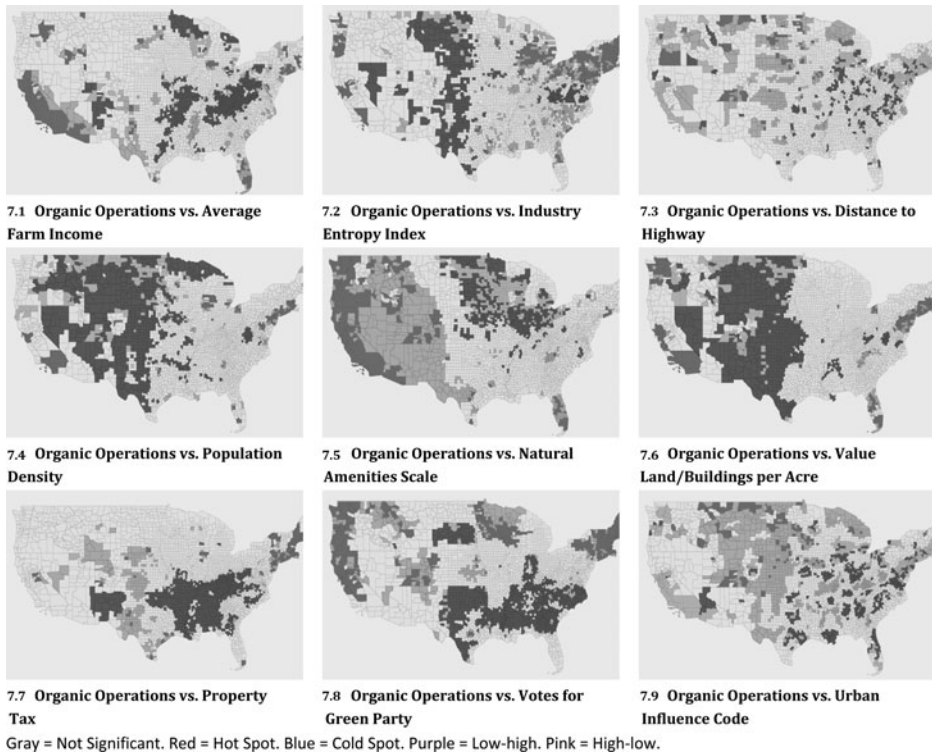


Figure 7. All Organic Operations versus County-level Characteristics

that population density, farm income, the presence of natural amenities, land values, proximity to urban areas, and support for the Green Party are significantly associated with the presence of hot spots.

To better place our findings within the context of existing literature and of future research, Figure 7 shows the relationships between the spatial distribution of organic operations and nine county-level variables that influence growth. Panels 7.1, 7.2, 7.4, 7.6, and 7.8 display many high-to-high and low-to-low areas and relatively few low-to-high and high-to-low areas, which indicates that organic operations tend to cluster in areas with a high average farm income, industry-entropy index, population density, value of land and buildings per acre, and receptiveness to organic (proxied by votes for the Green Party), respectively. On the other hand, panels 7.3 and 7.9 display many low-high and high-low areas and relatively few high-high and low-low areas, which suggests that organic operations tend to cluster in areas with shorter distances to highways and lower urban influence codes, respectively. The pattern is less clear for the score on the natural amenities scale (7.5) and property taxes (7.7). Further research is required to fully capture the effect of county-level characteristics on the spatial distribution of organic operations.

Methodology and Data: Identification of Spatial Autocorrelation by Type

In the second part of our analysis, we identify types of spatial autocorrelation present in the distribution of certified organic operations. In regressions, spatial autocorrelation typically takes two forms: spatial lag and spatial error. To illustrate this concept, we begin with the general spatial autoregressive model (LeSage 1998, Anselin 1999):

$$y = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{x} \beta + \mu$$

where $\mu = \lambda \mathbf{W}_2 \mu + \varepsilon$, ε is distributed $N(0, \sigma^2 I_n)$, and \mathbf{W}_1 and \mathbf{W}_2 are weighting matrices. As explained in Schmidtner et al. (2012), the significance of λ implies that some spatially correlated explanatory variables are important in the error specification (spatial error/heterogeneity) and the significance of ρ implies that some explanatory variables that are correlated with the average of the dependent variable are important in the model specification (spatial lag).

In this case, \mathbf{y} is a matrix of elements representing the number of certified organic operations as a share of total establishments in a particular county.⁸ The elements in the \mathbf{x} matrix represent county-level variables related to economic, demographic, climate, land, and political characteristics. To provide additional insight and check the robustness of our results, we also define alternative versions of \mathbf{y} in which the elements represent the share of certified organic operations that have a specific primary scope (crops, handling, and livestock).⁹

For the weighting matrices \mathbf{W}_1 and \mathbf{W}_2 , we use a row-standardized inverse-distance weighting matrix as we believe that is the most appropriate specification for our data. A common alternative is the contiguity-based weighting matrix, which would typically assign a weight of 1 to counties that are either adjacent to each other or have a shared corner and a weight of 0 otherwise (LeSage 1998, Anselin 1999). In our case, however, missing county-level data causes some observations to be dropped and a contiguity-based weighting matrix would create “islands” that do not exist. Distance-band weighting is another common alternative, but a brief analysis suggested that its results would not be significantly different from those of the inverse-distance weighting matrix. First, when comparing these two matrices, we find that they are correlated with a correlation coefficient of 0.6469. We also find that the correlation coefficient between the spatially lagged dependent variable with the inverse-distance matrix and the spatially lagged dependent variable with the distance-band matrix is 0.8688 and that the correlation coefficient between the spatially lagged matrix of independent variables with

⁸ For comparison and for consistency with our cluster identification, we also compare results using the count, rather than the share, of organic operations as the dependent variable.

⁹ Again, we also compare the counts.

the inverse-distance matrix and the spatially lagged matrix of independent variables with the distance-band matrix is 0.7548. These results are consistent with LeSage and Pace (2010), which noted that estimates from spatial models are not highly sensitive to the specification of the weighting matrix.

Although simple statistical tests are available for spatial lag and spatial error in a linear regression, we need to consider potential nonlinearity of our model. Close to half of the counties in our data set do not have any certified organic operations so our dependent variable is censored at 0. To account for this censoring, we consider a Tobit model (Sigelman and Zeng 1999) instead of a linear model when testing for spatial lags and errors. The Tobit model is defined as

$$y_i^* = x_i' \beta + \varepsilon_i$$

where y_i^* is an unobserved latent variable (Greene 2008). The observed variable, y_i , is determined as follows.

$$y_i = 0 \text{ if } y_i^* \leq 0$$

$$y_i = y_i^* \text{ if } y_i^* > 0$$

Because we are aware of no established spatial autocorrelation tests for Tobit models, we estimate a Tobit spatial autoregressive model implemented in Matlab to determine whether our model exhibits a statistically significant spatial lag and to examine the statistical significance of the coefficient of the spatially lagged dependent variable. To test for the presence of spatial error, we estimate a Tobit spatial error model and look at the statistical significance of the coefficient for the spatially lagged error term (LeSage 2001).

When choosing an estimator for our spatial Tobit models, we need to take several details into account. First, according to literature addressing spatial autoregressive models (e.g., LeSage 1998, Kelejian and Prucha 1998, Fingleton and Le Gallo 2008), the spatial lag variable is endogenous because it captures simultaneous spatial interactions (Fingleton and Le Gallo 2008). Second, the estimator must account for the multivariate normal distribution of the dependent variable (LeSage and Pace 2009).¹⁰ We therefore use a Gibbs sampling method to estimate a Bayesian spatial Tobit heteroskedastic model. This method appropriately defines the distributions of all of the parameters of interest and allows for sampling from multivariate and unknown distributions. The distribution of a parameter value of interest is calculated using an algorithm that starts with the initial values of the parameter and updates those values based on the other parameters and on draws from the appropriate distribution (formulas provided in LeSage (1998,

¹⁰ We thank James LeSage for this note.

1999b)). To implement this method, we use LeSage's (2001) spatial econometric toolbox. To assess the accuracy and convergence of the Gibbs sampler, we follow LeSage (1999b) and examine estimates of autocorrelation, Rafterty-Lewis Markov-chain Monte Carlo convergence diagnostics, and Geweke numerical standard errors, relative numerical efficiency estimates, and chi-squared tests.

The data on certified organic operations for this analysis again come from USDA's National Organic Program. Data for the county-level variables come from the U.S. Census Bureau (2010) and USDA's (2007) Census of Agriculture.

Economic theory provides little guidance regarding specific variables to include in \mathbf{x} , which controls for factors other than the number of neighboring organic operations that may affect the spatial distribution of the organic operations, and we rely extensively on the current empirical literature to determine its specification. Table 1 provides descriptions and descriptive statistics for the variables chosen, and Table 2 presents our rationale for including each variable, assigns the variables to one of eight categories (opportunity cost, resources, risk, size, urban sprawl, receptiveness to organic, market access, and time spent on the farm), and cites the studies on which inclusion of the variables were based. We assigned the variables to categories solely to provide justification for their inclusion. We are mainly interested in testing for the presence of spatially lagged dependent variables and/or errors and do not attempt to test any hypotheses regarding the categories in which the independent variables fit. Because some of the studies of organic operations and clustering have yielded mixed results, variables sometimes fit into more than one category. For example, a negative coefficient for land values would fit into opportunity costs while a positive coefficient for the same variable would fit into resources.

To mitigate potential simultaneity and/or endogeneity, we add time lags for many of the variables in \mathbf{x} . Specifically, we use data from 2009 for \mathbf{y} and data from 2007 for many of the variables included in \mathbf{x} , especially those that represent economic conditions and could affect the location of organic operations if concurrent data are used. Any other unaccounted-for factors are captured by the spatial error term, which captures spatial autocorrelation in the residuals.

Results of the Analysis of Spatial Autocorrelation

As shown in Table 3, all of the spatial autoregressive models (share of all organic operations and share of organic operations by primary scope) pass the accuracy and convergence tests and point to the presence of both spatial lags and spatial error¹¹ (positive and significant spatial lag and spatial error terms). The positive spatial lag implies that organic operations in general and

¹¹ This also holds for models that used the corresponding count variables.

Table 1. Summary Statistics and Description

Variable	Description	Obs.	Mean	Std. Dev.	Min.	Max.	Source
<i>Organic share</i>	Number of certified organic operations divided by the total number of establishments (2009)	2,031	0.0091982	0.0298247	0	0.983871 ^a	National Organic Program and U.S. Census
<i>Organic number</i>	Number of certified organic operations (2009)	2,031	7.358444	18.97139	0	287	National Organic Program
<i>Org share-crops</i>	Number of certified organic operations for which primary scope is crops divided by the total number of establishments (2009)	1,958	0.005971	0.020225	0	0.596774	National Organic Program and U.S. Census
<i>Org share-handler</i>	Number of certified organic operations for which primary scope is handling divided by the total number of establishments (2009)	1,958	0.00233	0.0112	0	0.387097	National Organic Program and U.S. Census

<i>Org share-live</i>	Number of certified organic operations for which primary scope is livestock divided by the total number of establishments (2009)	1,958	0.000617	0.003336	0	0.08284	National Organic Program and U.S. Census
<i>Org number-crops</i>	Number of certified organic operations for which primary scope is crops (2009)	1,958	4.42288	12.70936	0	227	National Organic Program
<i>Org number-handler</i>	Number of certified organic operations for which primary scope is handling (2009)	1,958	1.897344	5.92727	0	133	National Organic Program
<i>Organic number-live</i>	Number of certified organic operations for which primary scope is livestock (2009)	1,958	0.602656	2.795615	0	44	National Organic Program
<i>Retail sales-07</i>	Retail sales for human consumption in 2007 in dollars	2,031	447.389	1051.356	1	17,170	Census of Agriculture
<i>Ag receipts-07</i>	Receipts from agricultural services and custom work in 2007 in dollars	2,031	771.1359	1469.303	1	28,290	Census of Agriculture

Continued

Table 1. Continued

Variable	Description	Obs.	Mean	Std. Dev.	Min.	Max.	Source
<i>Temp workers-07</i>	Number of workers hired for less than 150 days, 2007	2,031	686.9581	1988.774	38	52,428	Census of Agriculture
<i>Conserve receipts-07</i>	Receipts from federal conservation and wetland programs measured in dollars per operation, 2007	2,031	4.222664	4.326851	0.067	50.534	Census of Agriculture
<i>Commodity sales-07</i>	Total commodity sales in 2007 in dollars	2,031	4074.684	7412.908	20	160,498	Census of Agriculture
<i>Avg farm income-07</i>	Receipts of income and farm-related totals measured in dollars per operation, 2007	2,031	14.79669	10.19674	2.001	100.165	Census of Agriculture
<i>Avg commute time-07-11</i>	Average travel time to work for workers 16 years and older not working at home, 2007–2011	2,031	22.58203	5.072798	8.5	42.5	U.S. Census
<i>Pop density-09</i>	Population per square mile in 2009	2,031	129.9401	264.9883	0.26522	3,002.117	U.S. Census
<i>Nat amenities scale-09</i>	Natural amenities scale	2,031	-0.37106	2.109244	-6.4	11.17	USDA Economic Research Service
<i>Land values-07</i>	Value of land and buildings per acre, 2007	2,031	3.004897	2.3005	0	27.122	Census of Agriculture
<i>Dist to interstate-07</i>	Distance of the county from an interstate highway measured in kilometers	2,031	12.12979	22.28624	0	148.0272	U.S. Geological Survey (distance calculated in GIS)

<i>Urban areas</i>	Number of urban areas/ clusters in the county	2,031	1.770556	1.668172	0	15	U.S. Census (number calculated in GIS)
<i>Politics</i>	Dummy variable for the county voting for Obama in the 2012 election	2,031	0.194485	0.395902	0	1	Politico
<i>Insurance numbers-07</i>	Number of operations participating in crop insurance programs, 2007	2,031	149.1467	163.1179	1	2,115	Census of Agriculture

^a The maximum value corresponds with Kidder County, North Dakota, which is a sparsely populated, rural county (U.S. Census Bureau 2010). The second highest share is 0.341615.

Table 2. Independent Variables and Prior Studies by Category

Category	Variable	Direction of Correlation Based on Literature	Source
Opportunity cost: A high value of assets implies a high opportunity cost of using them for farming.	<i>Land values-07</i>	Negative: Higher value of land and buildings per acre implies higher opportunity costs of farming.	Brown, Goetz, and Fleming (2012), Mishra and Goodwin (1997)
	<i>Nat amenities scale-09</i>	Negative: A higher natural amenities scale implies a higher value of the location as a place to live, which may imply a higher opportunity cost.	
Resources: Availability of labor and capital affects farm viability. Organic operations tend to rely more on their own resources.	<i>Ag receipts-07</i>	Ambiguous: While Argiles and Brown (2010) and Schmidtner et al. (2012) mentioned that organic farms may rely more on their own resources than conventional farms, Brown et al. (2012) suggested that the availability of agricultural services may help farms that are unable to afford to buy their own harvesting equipment. Since organic farms tend to be smaller, they may be less likely to afford machinery. It also may be more efficient for them not to buy their own machinery given their small scale.	Argiles and Brown (2010), Schmidtner et al. (2012), Brown, Goetz, and Fleming (2012), Lotter (2003), Klonsky and Tourte, (1998), Goetz (1997)
	<i>Pop density-09</i>	Positive: This variable may capture labor supply, which would imply a positive effect given the fact that organic operations tend to be more labor intensive.	
	<i>Land values-07</i>	Positive: A higher value of land and buildings per acre may imply availability of better resources.	
	<i>Nat amenities scale-09</i>	Positive: A higher natural amenities scale may imply availability of better resources.	

<p>Risk: Activities and programs that decrease the risk of farming (i.e., uncertain crop yields) and improve farm viability; however, the presence of these programs may also reflect the need for them.</p>	<p><i>Insurance numbers-07</i></p>	<p>Ambiguous: Crop insurance programs decrease the risk of losses from uncertain crop yields but may also imply unfavorable conditions that lead to the need for them.</p>	<p>Brown, Goetz, and Fleming (2012)</p>
<p>Size: Organic farms tend to be smaller and make less money on sales (75 percent of certified organic farms are under 2.5 hectares with about half grossing less than \$15,000).</p>	<p><i>Commodity sales-07/ Avg farm income-07/ Retail sales-07</i></p>	<p>Negative: Organic operations are smaller and associated with lower sales.</p>	<p>Lotter (2003), Klonsky and Tourte, (1998), Dimitri (2010)</p>
<p>Urban sprawl: Protection from sprawling development may be beneficial to farm viability.</p>	<p><i>Dist to interstate-07</i></p>	<p>Positive: A greater distance from a highway may imply protection from sprawling development.</p>	<p>Brown, Goetz, and Fleming. (2012)</p>
	<p><i>Urban areas</i></p>	<p>Negative: Being close to fewer urban areas may imply protection from sprawling development.</p>	
	<p><i>Pop density-09</i></p>	<p>Negative: Lower population density may imply protection from sprawling development.</p>	
	<p><i>Avg commute time-07-11</i></p>	<p>Positive: A longer travel time to work may imply protection from sprawling development.</p>	

Continued

Table 2. Continued

Category	Variable	Direction of Correlation Based on Literature	Source
Receptiveness to organic: Certain areas may be more receptive to organic than others.	<i>Conserve receipts-07</i>	Positive: High receipts from government conservation programs and wetlands may imply a socioeconomic environment that is receptive to the idea of organic agriculture.	Schmidtner et al. (2012)
	<i>Politics</i>	Positive: Liberal areas are generally associated with more receptiveness to organic.	
Market access: Proximity to roads and urban centers provide market access.	<i>Dist to interstate-07</i>	Negative: Being closer to a highway may provide market access, which encourages the development of farms.	Mishra, El-Osta, and Sandretto (2002), Brown, Goetz, and Fleming (2012)
	<i>Urban areas</i>	Positive: Being close to urban areas may provide market access, which encourages the development of farms.	
	<i>Pop density-09</i>	Positive: Being in highly populated areas may provide market access, which encourages the development of farms.	
	<i>Avg commute time-07-11</i>	Negative: A longer travel time to work may imply limited market access. As long as the population density is constant (i.e., nonsignificant), travel time to work can capture the distance travelled.	
Time spent on the farm: Organic farmers may be inclined to participate in off-farm activities.	<i>Commodity sales-07/ Avg farm income-07/ Retail sales-07</i>	Positive: Higher sales may imply greater market access.	Lotter (2003)
	<i>Temp workers-07</i>	Positive: For about two-thirds of organic farms, sales of organic products account for less than half of the net family income. This may indicate that organic farm owners work off-farm jobs (seasonality may also be an explanation).	

Table 3. Tests for Spatial Lag and Spatial Error in the Tobit Models

Dependent Variable	Tobit Spatial Error Model	Tobit Spatial Autoregressive Model
	λ	ρ
<i>Organic share</i>	0.990727***	0.914124**
<i>Organic number</i>	0.982933***	0.963783**
<i>Org share-crops</i>	0.994286***	0.911364**
<i>Org number-crops</i>	0.972535***	0.957396**
<i>Org share-handler</i>	0.983324***	0.845932**
<i>Org number-handler</i>	0.936908 ***	0.924824**
<i>Org share-live</i>	0.970087***	0.952625**
<i>Organic number-live</i>	0.984535 ***	0.977669**

Note: * Significant at the 10 percent level. ** Significant at the 5 percent level. *** Significant at the 1 percent level.

when broken down by primary scope tend to cluster near areas that have a high number of organic operations. The significant spatial error implies that variables that are not included in our model likely affect the spatial distribution of certified organic operations.

Since we find that both spatial lag and spatial error are present, we cannot determine whether the spatial autoregressive or spatial error model is most appropriate. To our knowledge, there is no established method by which to compare the appropriateness of these two models, but LeSage and Pace (2009) suggested an interesting solution: a spatial Durbin model is the proper specification when the probabilities of the correct model specification being a spatial autoregressive model and of being a spatial error model are both greater than zero. A spatial Durbin model (LeSage 1998) is defined as

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{x}\beta_1 + \mathbf{W}\mathbf{x}\beta_2 + \varepsilon.$$

According to LeSage (1998), the results of this model take the possibility of spatial lag in both the dependent variable and in the independent variables into account. In our case, the coefficients in front of the spatially lagged independent variables show the effects of changes in the values of the independent variables in one county on the share of organic operations in neighboring counties. We therefore estimate spatial Durbin Tobit models.

The results of the spatial Durbin Tobit models, all of which pass the accuracy and convergence tests, suggest the presence of spatial autocorrelation and spillovers when analyzing the shares and counts of all organic operations and of the three types of organic operations (crops, handling, and livestock). Table 4 presents the results of the spatial Durbin Tobit model, which

Table 4. Results: Share of Operation That Is Organic

Variable	Coefficient	Interpretation/Fit with Literature	Direct Effects	Indirect Effects
Constant	-127.663204**			
Retail sales-07	-0.000043		-0.000048	
Ag receipts-07	0.000129*	Resources (beneficial to organic)	0.000145	
Temp workers-07	0.000039		0.000037	
Conserve receipts-07	0.057305**	Receptiveness	0.058739	
Commodity sales-07	-0.000002		-0.000002	
Avg farm income-07	-0.010543		-0.011984	
Avg commute time-07-11	0.037465**	Urban sprawl	0.03457	
Pop density-09	-0.000259		-0.000243	
Nat amenities scale-09	0.037212		0.03303	
Land values-07	-0.041379		-0.03714	
Dist to interstate-07	0.007262**	Urban sprawl	0.007521	
Urban areas	-0.182286**	Urban sprawl	-0.186243	
Politics	-0.02731		0.029495	
Insurance numbers-07	0.000586		0.000459	
W-Retail sales-07	-0.001506			-0.00904
W-Ag receipts-07	0.004752**	Resources (beneficial to organic)		0.028536
W-Temp workers-07	-0.000553			-0.00305
W-Conserve receipts-07	0.388469			2.555781
W-Commodity sales-07	0.000029			0.000155
W-Avg farm income-07	-0.42524**	Size		-2.54931
W-Avg commute time-07-11	-0.900275**	Market access		-5.11773

<i>W</i> -Pop density-09	0.004861		0.026724
<i>W</i> -Nat amenities scale-09	-1.283749*	Opportunity cost	-7.34866
<i>W</i> -Land values-07	1.308321*	Resources	7.520867
<i>W</i> -Dist to interstate-07	0.071512		0.456497
<i>W</i> -Urban areas	-1.045417		-7.00379
<i>W</i> -Politics	17.145372**	Receptiveness	100.5659
<i>W</i> -Insurance numbers-07	-0.038746**	Risk (need)	-0.2248
ρ	0.693221**	Similar observations tend to be close together	
Number of observations	2,031		
Number censored	663		

Notes: A Bayesian spatial autoregressive Tobit heteroskedastic model yields similar estimates but with higher magnitudes and more significant variables.
 *Significant at the 10 percent level. ** Significant at the 5 percent level. *** Significant at the 1 percent level.

analyzes the share of all organic operations regressed against county-level factors. The positive significant coefficient ρ for the spatially lagged dependent variable implies that similar observations tend to be located close together; that is, counties that have a large share of organic operations tend to be relatively close to other counties that have large shares of organic operations.¹² As with hot spots, this result indicates that the intensity of organic operations of neighboring counties is positively related to a county's intensity of organic operations.

Seven of the fourteen spatially lagged independent variables are significant, suggesting the presence of cross-county spillover. For example, the positive significant coefficient for agricultural receipts (*W-Ag receipts-07*) indicates that availability of agricultural services in a county is positively related to the intensity of organic operations in neighboring counties. Simultaneously, there is a significant positive coefficient for land values (*W-Land values-07*) and a significant negative coefficient for the natural amenity score (*W-Nat amenities scale-09*). Considered together, these results point to cross-county spillovers for inputs and farm resources. The positive significant coefficient for support for the Democratic Party (*W-Politics*) suggests that political receptiveness to organic in one county may spill over into neighboring counties. Similarly, the negative coefficient for the number of organic operations participating in crop insurance programs (*W-Insurance numbers-07*) suggests that perceptions of risk associated with farming may also spill over into neighboring counties. The negative significant coefficient for average farm income (*W-Avg farm income-07*) indicates that the intensity of organic operations in a county is relatively high when the county is surrounded by counties that have a large number of small farms. The negative significant coefficient on travel time to work (*W-traveltimetowork-07-11*) suggests that a county's level of market access spills over to neighboring counties.

It is also interesting to examine the direct and indirect effects of the independent variables. The indirect effect of a variable represents the sum of its effects on observations outside of its area ($\sum_{(j \neq i)} \partial y_j / \partial \mathbf{W}x_i$), and the direct effect represents the variable's effect on observations inside its area ($\partial y_i / \partial x_i$). The total effect is the sum of the direct and indirect effects (LeSage and Pace 2009). We find that the directions of the direct effects are consistent with the coefficients of the independent variables with no spatial lags and that the directions of the indirect effects are consistent with the coefficients of the spatially lagged independent variables.

Table 5 presents the results of the same analysis applied to the share of organic operations broken down by primary scope. As with the results for organic operations in general (crops, handling, and livestock), the coefficient

¹² The model that used the corresponding count variable also yielded a positive and significant spatially lagged dependent variable but the coefficients for the independent variables were somewhat different.

Table 5. Results of Estimation by Primary Organic Scope

Variable	Crops	Effect	Handling	Effect	Livestock	Effect
Direct Effects						
Constant	0.044496*		0.055419*		0.098823**	
Retail sales-07	0.00583	0.001893	0.006035	0.005523	-0.018106	-0.00348
Ag receipts-07	0.060198*	0.062538	0.072958**	0.0698	0.029281	0.022195
Temp workers-07	0.026485	0.025074	0.00657	0.007345	0.004592	0.003307
Conserve receipts-07	0.06767**	0.071405	0.047824*	0.049554	0.010455	0.006418
Commodity sales-07	-0.00091	-0.00037	-0.015087	-0.01372	0.045043*	0.042427
Avg farm income-07	-0.053123**	-0.05967	-0.032569	-0.03378	-0.031444	-0.03842
Avg commute time-07-11	0.081378**	0.075737	0.059173**	0.05749	0.044239*	0.03241
Pop density-09	-0.00782	-0.00469	-0.007243	-0.00575	-0.008268	-0.01282
Nat amenities scale-09	0.005255	0.003776	-0.036111	-0.03713	-0.041537	-0.04025
Land values-07	-0.05289	-0.04875	-0.010392	-0.01005	-0.017697	-0.02446
Dist to interstate-07	0.034348	0.038332	0.021142	0.023483	0.089441**	0.083208
Urban areas	-0.103593**	-0.1042	-0.047011*	-0.04652	-0.050557*	-0.05568
Politics	0.00356	0.012367	0.040845*	0.041975	-0.016884	-0.00281
Insurance numbers-07	0.02975	0.024404	0.005739	0.00586	0.005325	0.001885
Indirect Effects						
W-Retail sales-07	-1.179052**	-7.40902	-0.367035	-1.10129	2.305392**	27.58374
W-Ag receipts-07	0.642246	4.367561	-2.224181**	-6.54877	-1.137373	-13.3984
W-Temp workers-07	-0.44652	-2.6732	0.51674	1.564406	-0.208365	-2.45848
W-Conserve receipts-07	1.052494**	6.987673	1.149518**	3.560505	-0.643363*	-7.63358

Continued

Table 5. Continued

Variable	Crops	Effect	Handling	Effect	Livestock	Effect
<i>W-Commodity sales-07</i>	0.165886	1.035439	0.958572	2.852855	-0.448757	-4.91476
<i>W-Avg farm income-07</i>	-1.906339**	-12.3028	-0.813779*	-2.53375	-1.063943**	-13.2269
<i>W-Avg commute time-07-11</i>	-1.736015**	-10.5253	-1.178692**	-3.44388	-1.885721**	-22.3148
<i>W-Pop density-09</i>	0.940148	5.859772	1.035576	3.093935	-0.705305	-8.60007
<i>W-Nat amenities scale-09</i>	-0.43847	-2.71686	-0.655695	-2.03254	0.242238	2.480588
<i>W-Land values-07</i>	1.28513**	7.850575	0.271042	0.8185	-1.040558**	-12.6909
<i>W-Dist to interstate-07</i>	1.16009*	7.500384	1.602668**	4.878192	-1.057128	-11.7347
<i>W-Urban areas</i>	-0.09388	-1.11196	0.372533	1.041539	-0.756047	-9.64913
<i>W-Politics</i>	2.617899**	16.47569	0.744992*	2.316695	2.219676**	26.57333
<i>W-Insurance numbers-07</i>	-1.618736**	-10.0474	0.065034	0.204524	-0.544982	-6.53967
ρ	0.726291**		0.468207*		0.860965**	
Number of observations	1,875		1,875		1,875	
Number censored	1,441		1,509		1,667	

Notes: * Significant at the 10 percent level. ** Significant at the 5 percent level. *** Significant at the 1 percent level.

for the spatially lagged dependent variable is positive and significant. Handling has the lowest coefficient and livestock has the highest.¹³

There are some differences, however, in the spatially lagged independent variables. For crop operations, the coefficients are negative and significant for average farm income, average commute time, and number of crop insurance participants and positive and significant for land values and political support for the Democratic Party. The coefficient for distance to the nearest interstate is positive and significant, indicating that the negative effects of urban sprawl may spill over county lines, and the positive significant coefficient for receipts from conservation programs (*W-Conserve receipts-07*) implies that the effects of receptiveness to organic may spill over county lines. The negative significant coefficient for retail sales suggests that the concentration of organic operations in a county is higher when the county is surrounded by counties with large numbers of small establishments.

The results for the share of organic operations with handling as the main scope are similar, except for the lack of significance for retail sales, land values, and number of participants in crop insurance programs and a positive significant coefficient for agricultural receipts. In the results for the share of organic operations with livestock as the main scope, the coefficient for support for the Democratic Party is significant and positive and the coefficients for average farm income and travel time to work are significant and negative, which is consistent with the results for crops and handling. However, the positive significant coefficient for retail sales (which fits into the market access category) and the negative significant coefficients for receipts from conservation programs (which fits into the risk category) and land values (which fits into the opportunity cost category) are not consistent with the results of the other groups.

The coefficients for the independent variables without spatial lags are less significant and are not necessarily consistent with the coefficients for the corresponding spatially lagged variables. Table 4 shows that the coefficients are significant for only five of the variables. The positive significant coefficient for agricultural receipts fits into the resource category and is in line with results of prior studies that showed that greater availability of agricultural services can be beneficial to organic farms. The positive significant coefficient on receipts from conservation programs fits into the category of receptiveness to organic, supporting the notion that organic operations are likely to be more prevalent in areas that are receptive to the idea of organic production. The positive significant coefficients for average commute time and distance to the nearest interstate and the negative

¹³ The models that used the count of organic livestock and crop operations also yielded positive and significant spatially lagged dependent variables but the coefficients for the independent variables were somewhat different. The model that used the count of handling operations yielded a statistically insignificant spatially lagged dependent variable.

coefficient for the number of urban areas all fit into the category of urban sprawl, in which organic farms tend to fare better when they are sheltered from urban development.

Table 5 shows that the results for the share of organic operations with crops and handling as the main scope are similar to those in Table 4, with the exception of the nonsignificant coefficient for distance to the nearest interstate and the negative significant coefficient for average farm income (crops only), which fits into the size category. For share of livestock operations, we find that the coefficient for the number of urban areas is negative and significant and the coefficients for average commute time and distance to the nearest interstate are positive and significant. The coefficient for commodity sales is also positive and significant, placing it in the market access category.

Conclusion

We document statistically significant hot spots and cold spots of certified organic operations under a broad definition and compare those results to organic hot spots and cold spots under three narrower definitions—crop producers, livestock producers, and handlers. We then compare all of those maps of distributions of organic hot spots and cold spots to those of agricultural farm establishments generally and of all business establishments.

We find that the largest hot spots for organic operations in general, crop operations, and handling operations occur along the West Coast of the United States from southern California to northern Washington. This extensive organic hot spot closely matches one of the hot spots for agricultural operations in general. However, many of the other organic hot spots do not match hot spots for general agriculture. This is particularly true for the Northeast, where we find numerous organic hot spots and almost no agricultural hot spots, and for the South and Southeast, where we find numerous agricultural hot spots and organic cold spots. Our spatial analysis thus demonstrates that organic operations follow a different geographic concentration pattern than agricultural operations generally and business establishments generally.

We also confirm the presence of spatial autocorrelation using various spatial autoregressive models that examine the shares or counts of certified organic operations. Spatial dependence is confirmed for a large number of variations in how organic operations are measured and for a number of spatial econometric models.

The results of this study clearly demonstrate the importance of spatial spillovers in organic agriculture. As a matter of policy, spatial dependence can be an important consideration for USDA and state governmental agencies that provide assistance to organic farmers. Our results suggest, for example, that public and private policies aimed at promoting organic agriculture might be most effective in areas where organic agriculture already has a foothold.

In the future, a spatio-temporal analysis of organic operations over time could provide a fuller picture of the areas in the United States in which organic agriculture is prevalent and growing.

Our analysis of factors associated with the distribution of organic operations provides a starting point for further exploration of the formation of hot spots of organic operations. The results suggest, for example, that the degree of urban influence and distance to highways are negatively related to the formation of organic hot spots. We also find that the three types of organic operations—crop producers, livestock producers, and handlers—respond differently to these factors, potentially providing insight into their functions. One potentially interesting avenue for further study would be to analyze why receipts from agricultural services and custom work appear to have a significant association with organic crops and handling operations but not with organic livestock and why the distance to a highway appears to have a significant effect on organic livestock operations but not on crop and handling operations.

Our research also could be useful for understanding the consequences of the rapidly expanding demand for organic products and the potential for supply shortages. We find that areas of the United States known to be agricultural strongholds, such as the corn and soybean belts, are not hot spots of organic production. This research implicitly shows that some barriers exist in extending agricultural success in these prevalent and highly valued crops to similar success in their organic counterparts, which are known to be in great demand. Our results potentially provide a logical first step to a deeper investigation into organic agriculture's role as a rural development tool. Documenting hot spots and cold spots is a necessary starting point for investigating their potential economic impact on local economies.

References

- Anselin, L. 1995. "Local Indicators of Spatial Association - LISA." *Geographical Analysis* 27(2): 93-115.
- . 1999. "Spatial Econometrics." Center for Spatially Integrated Social Science, University of California, Santa Barbara.
- Anselin, L., I. Syabri, and Y. Kho. 2006. "GeoDa: An Introduction to Spatial Data Analysis." *Geographical Analysis* 38(1): 5-22.
- ArcGIS Resource Center. 2012. "How Cluster and Outlier Analysis (Anselin Local Moran's I) Works." http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_Cluster_and_Outlier_Analysis_Anselin_Local_Moran_s_I_works/005p00000012000000 (accessed August 10, 2012).
- Argiles, J.M., and N.D. Brown. 2010. "A Comparison of the Economic and Environmental Performances of Conventional and Organic Farming: Evidence from Financial Statements." *Agricultural Economics Review, Greek Association of Agricultural Economists* 11(1): 69.
- Banasick, S., G. Lin, and R. Hanham. 2009. "Deviance Residual Moran's I Test and Its Application to Spatial Clusters of Small Manufacturing Firms in Japan." *International Regional Science Review* 32(1): 3-18.

- Barkley, D., and M. Henry. 1997. "Rural Industrial Development: To Cluster or Not to Cluster." *Review of Agricultural Economics* 19(2): 308–325.
- Brown, J.P., S.J. Goetz, and D.A. Fleming. 2012. "Multifunctional Agriculture and Farm Viability in the United States." Presentation, Agricultural and Applied Economics Association 2012 annual meeting, Seattle, WA.
- Cainelli, G. 2008. "Spatial Agglomeration, Technological Innovations, and Firm Productivity: Evidence from Italian Business Districts." *Growth and Change* 39(3): 414–435.
- California Certified Organic Farmers. Primary website, www.ccof.org (accessed November 22, 2012).
- Chevassus-Lozza, E., and D. Galliano. 2003. "Local Spillovers, Firm Organization, and Export Behaviour: Evidence from the French Food Industry." *Regional Studies* 37(2): 147–158.
- Dimitri, C. 2010. "Organic Agriculture: An Agrarian or Industrial Revolution?" *Agricultural and Resource Economics Review* 39(3): 384.
- Dimitri, C., and L. Oberholtzer. 2009. "Marketing U.S. Organic Foods: Recent Trends from Farms to Customers." Economic Information Bulletin EIB-58, Economic Research Service, USDA, Washington, DC. Available at www.ers.usda.gov/publications/eib-economic-information-bulletin/eib58.aspx (accessed September 26, 2012).
- Duranton, G., and D. Puga. 2004. "Micro-foundations of Urban Agglomeration Economies." *Handbook of Regional and Urban Economics* 4(2004): 2063–2117.
- Eades, D., and C. Brown. 2006. "Identifying Spatial Clusters within U.S. Organic Agriculture." Research Paper 2006–10, Regional Research Institute, West Virginia University.
- Economic Research Service. 2008. "U.S. Census of Agriculture." ERS, U.S. Department of Agriculture, Washington, DC. Available at www.agcensus.usda.gov/Publications/2007/Online_Highlights/Organics/index.php (accessed October 21, 2012).
- . 2012. "Higher Food Prices and Local Food Labels." ERS, U.S. Department of Agriculture, Washington, DC. Available at www.ers.usda.gov/media/452879/eib55c_1.pdf (accessed November 22, 2012).
- . 2013. Natural Amenities Scale, www.ers.usda.gov/data-products/natural-amenities-scale.aspx (accessed January, 2013)
- Environmental Protection Agency. 2014. "Ag 101" web page, www.epa.gov/oecaagct/ag101/cropmajor.html (accessed October 10, 2014). Environmental Protection Agency, Washington, DC.
- Fairfield, H. 2009. "The Hot Spots for Organic Food." *The New York Times*, Business Section, May 2. Available at www.nytimes.com/2009/05/03/business/03metrics.html?_r=0 (accessed June 11, 2015).
- Farmer, J.R. 2014. "Organic Farming in West Virginia: A Behavioral Approach." *Journal of Agriculture, Food Systems, and Community Development* 4(4): 155–171.
- Farmer, J., C. Peters, E. Hansen, F. Boettner, and M. Betcher. 2013. "Overcoming the Market Barriers to Organic Production in West Virginia." School of Public Health, Indiana University, Bloomington, and Downstream Strategies, Morgantown, WV. Available at http://t.downstreamstrategies.com/documents/reports_publication/overcoming_market_barriers_to_organic_production_in_wv.pdf (accessed June 12, 2015).
- Feser, E.J. 1998. "Enterprises, External Economies, and Economic Development." *Journal of Planning Literature* 12(3): 283–302.
- Feser, E.J., and E.M. Bergman. 2000. "National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis." *Regional Studies* 34(1): 1–19.
- Feser, E., H. Renski, and H. Goldstein. 2008. "Clusters and Economic Development Outcomes: An Analysis of the Link between Clustering and Industry Growth." *Economic Development Quarterly* 22(4): 324–344.
- Fingleton, B., and J. Le Gallo. 2008. "Estimating Spatial Models with Endogenous Variables: A Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties." *Papers in Regional Science* 87(3): 319–339.

- Gabe, T.M. 2004. "Industry Agglomeration and Investment in Rural Businesses." *Review of Agricultural Economics* 27(1): 89–103.
- Gabe, T. 2008. "Impact of Agglomerations on the Economy." In S.J. Goetz, S. Deller, and T. Harris, eds., *Targeting Regional Economic Development*. Abingdon, Oxon, UK: Routledge Publishing.
- GeoDa Center. 2013. Data, <https://geodacenter.asu.edu> (accessed May 8, 2013). GeoDa Center, Arizona State University, Tempe.
- Gibbs, R.M., and G.A. Bernat. 1997. "Rural Industry Clusters Raise Local Earnings." *Rural Development Perspectives* 12(3): 18–25.
- Glaeser, E.L., H.D. Kallal, J.A. Scheinkman, and A. Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100(6): 1126–1152.
- Goetz, S. 1997. "State and County Level Determinants of Food Manufacturing Establishment Growth: 1987–93." *American Journal of Agricultural Economics* 79(3): 838–850.
- Goetz, S.J., M. Shields, and Q. Wang. 2008. "Identifying Food Industry Clusters: A Comparison of Analytical Tools." In S.J. Goetz, S. Deller, and T. Harris, eds., *Targeting Regional Economic Development*. Abingdon, Oxon, UK: Routledge Publishing.
- Graham, D.J., and H.Y. Kim. 2008. "An Empirical Analytical Framework for Agglomeration Economies." *Annals of Regional Science* 42(2): 267–289.
- Greene, W.H. 2008. *Econometric Analysis* (sixth edition). Upper Saddle River, NJ: Pearson Prentice Hall.
- Greenstone, M., R. Hornbeck, and E. Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy* 118(3): 536–598.
- Grogan, K.A., and R.E. Goodhue. 2012. "Spatial Externalities of Pest Control Decisions in the California Citrus Industry." *Journal of Agricultural and Resource Economics* 37(1): 156–179.
- Hatzenbuehler, P.L., J.M. Gillespie, and C.E. O'Neil. 2012. "Does Healthy Food Cost More in Poor Neighborhoods? An Analysis of Retail Food Cost and Spatial Competition." *Agricultural and Resource Economics Review* 41(1): 43.
- Hooker, N., and C.J. Shanahan. 2012. "Emerging Spatial Dependencies within U.S. Organic Supply Chains." *Journal of Food Products Marketing* 18(5): 426–450.
- Hughner, R.S., P. McDonagh, A. Prothero, C.J. Shultz, and J. Stanton. 2007. "Who Are Organic Food Consumers? A Compilation and Review of Why People Purchase Organic Food." *Journal of Consumer Behavior* 6(2/3): 94–110.
- Jaenicke, E.C., S.J. Goetz, P-C. Wu, and C. Dimitri. 2009. "Identifying and Measuring the Effect of Firm Clusters among Certified Organic Processors and Handlers." Presented at the 2009 Agricultural and Applied Economics Association annual meeting, Milwaukee, WI. Available at <http://ageconsearch.umn.edu/handle/49205> (accessed September 18, 2012).
- Kelejian, H., and I. Prucha. 1998. "A Generalized Spatial Two-stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finance and Economics* 17(1): 99–121.
- Klonsky, K., and L. Tourte. 1998. "Organic Agricultural Production in the United States: Debates and Directions." *American Journal of Agricultural Economics* 80(5): 1119–1124.
- LeSage, J.P. 1998. "Spatial Econometrics." Working paper, Department of Economics, University of Toledo, OH. Available at www.spatial-econometrics.com/html/wbook.pdf (accessed August 10, 2012).
- . 1999b. "Applied Econometrics using MATLAB." Manuscript, Department of Economics, University of Toronto, Canada.
- . 2001. "Econometrics Toolbox" Matlab software.
- LeSage, J.P., and R.K. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Taylor and Francis.

- . 2010. "The Biggest Myth in Spatial Econometrics." Social Science Research Network, Rochester, NY. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1725503.
- Lian, J., H. Gong, X. Li, Y. Sun, W. Zhao, and L. Zhu. 2009. "The Analysis of Economic Spatial Characteristics of Beijing-Tianjin-Hebei Metropolitan Region Based on GIS." *Seventeenth International Conference on Geoinformatics*: 1–5.
- Lopez, R.A., A.M. Azzam, and C. Lirón-España. 2002. "Market Power and/or Efficiency: A Structural Approach." *Review of Industrial Organization* 20(2): 115–126.
- Lotter, D.W. 2003. "Organic Agriculture." *Journal of Sustainable Agriculture* 21(4): 59–128.
- Marasteanu, I.J., and E.C. Jaenicke. Forthcoming. "The Role of U.S. Organic Certifiers in Organic Hotspot Formation." *Renewable Agriculture and Food Systems*.
- Mishra, A.K., H.S. El-Osta, and C.L. Sandretto. 2002. "Factors Affecting Farm Enterprise Diversification." *Agricultural Finance Review* 64(2): 151–166.
- Mishra, A.K., and B.K. Goodwin. 1997. "Farm Income Variability and the Supply of Off-farm Labor." *American Journal of Agricultural Economics* 79(3): 880–887.
- Moons, E., T. Brijs, and G. Wets. 2008. "Hot Spot Analysis: Improving a Local Indicator of Spatial Association for Application in Traffic Safety." *Computational Science and Its Applications* 5072: 221–231.
- Morrison Paul, C.J., and D.S. Seigel. 1999. "Scale Economies and Industry Agglomeration Externalities: A Dynamic Cost Function Approach." *American Economic Review* 89(1): 272–290.
- Naik, G., and V. Nagadevara. 2010. "Spatial Clusters in Organic Farming: A Case Study of Pulses Cultivation in Karnataka." IIM Bangalore Research Paper 316, Social Science Research Network, Rochester, NY. Available at <http://ssrn.com/abstract=2124370>.
- National Organic Program. 2009. "2009 List of Certified USDA Organic Operations." National Organic Program, Agricultural Marketing Service, USDA, Washington, DC. Available at <http://apps.ams.usda.gov/nop> (accessed September 29, 2012).
- . 2015. "Welcome to the National Organic Program." National Organic Program, Agricultural Marketing Service, USDA, Washington, DC. Available at www.ams.usda.gov/AMSV1.0/NOP (accessed June 17, 2015).
- Onken, K.A., J.C. Bernard, and J.D. Pesek Jr. 2011. "Comparing Willingness to Pay for Organic, Natural, Locally Grown, and State Marketing Program Promoted Foods in the Mid-Atlantic Region." *Agricultural and Resource Economics Review* 40(1): 33.
- Osteen, C., J. Gottlieb, and U. Vasavada (eds.). 2012. "Agricultural Resources and Environmental Indicators, 2012. EIB-98." ERS, U.S. Department of Agriculture, Washington, DC.
- Politico. 2012. Map, "2012 Presidential Election." www.politico.com/2012-election/results/map/#/President/2012 (accessed January, 2016).
- Porter, M.E. 2003. "The Economic Performance of Regions." *Regional Studies* 37(6/7): 549–578.
- Richards, T.J., S.F. Hamilton, and P.M. Patterson. 2010. "Spatial Competition and Private Labels." *Journal of Agricultural and Resource Economics* 35(2): 183–208.
- Rocha, H.O., and R. Sternberg. 2005. "Entrepreneurship: The Role of Clusters—Theoretical Perspectives and Empirical Evidence from Germany." *Small Business Economics* 24(3): 267–292.
- Roe, B., E.G. Irwin, and J.S. Sharp. 2002. "Pigs in Space: Modeling the Spatial Structure of Hog Production in Traditional and Nontraditional Production Regions." *American Journal of Agricultural Economics* 84(2): 259–278.
- Schmidtner, E., C. Lippert, B. Engler, A.M. Haring, J. Aurbacher, and S. Dabbert. 2012. "Spatial Distribution of Organic Farming in Germany: Does Neighbourhood Matter?" *European Review of Agricultural Economics* 39(4): 1–23.
- Schmit, T.M., and J.S. Hall. 2013. "Implications of Agglomeration Economies and Market Access for Firm Growth in Food Manufacturing." *Agribusiness* 29(3): 306–324.

- Sigelman, L., and L. Zeng. 1999. "Analyzing Censored and Sample-selected Data with Tobit and Heckit Models." *Political Analysis* 8(2): 167–182.
- Stewart, L., D.M. Lambert, M.D. Wilcox, and B.C. English. 2009. "Tennessee Agriculture and Forestry Industry Clusters and Economic Performance, 2001–2006." *Journal of Agricultural and Resource Economics* 34(1): 172–195.
- Taus, A., Y. Ogneva-Himmelberger, and J. Rogan. 2013. "Conversion to Organic Farming in the Continental United States: A Geographically Weighted Regression Analysis." *The Professional Geographer* 65(1): 87–102.
- U.S. Census Bureau. 2010. "U.S. Census 2010." U.S. Census Bureau, Washington, DC. Available at www.census.gov/2010census/data.
- U.S. Department of Agriculture. 2007. "Census of Agriculture." USDA, Washington, DC. Available at www.agcensus.usda.gov/Publications/2007.
- U.S. Geological Survey. 2014. "U.S. Geologic Survey" website. www.usgs.gov (accessed January, 2014).
- Zhang, T., and G. Lin. 2008. "Identification of Local Clusters for Count Data: A Model-based Moran's I Test." *Journal of Applied Statistics* 35(3): 293–306.