Hog Production and Agglomeration Economies: The Case of U.S. State-Level Hog Production

*Gibson Nene1 and Karina Schoengold2

1Assistant Professor, Department of Economics, University of Minnesota Duluth.
2Associate Professor, Department of Agricultural Economics and School of Natural Resources, University of Nebraska-Lincoln.

The study examines the importance of agglomeration economies on U.S. hog production for the period, 1994-2006. Results suggest that hog production in a state is positively affected by hog production in a nearby state, confirming the presence of agglomeration economies at the state-level. This finding is true for both the top 22 hog producing states and for hog production in the Midwest region of the U.S. Agglomeration economies played an important role in the reorganization of the U.S. hog industry during the study period. In addition to agglomeration economies, our results also show that environmental regulations, hog price, land value, labor cost and the cost of corn input were important in shaping the U.S. hog industry during the study period.

Key words: U.S. hog industry, Agglomeration economies, Hog production, Geographical concentration, Midwest

JEL Classifications: L22, L25

INTRODUCTION

The U.S. hog production industry went through several changes and has become more geographically concentrated since the mid-1990s (Hubbell and Welsh, 1998; Herath, Weersink, and Carpentier, 2005a; Azzam, Nene and Schoengold, 2015). The industry which has experienced a huge decline in small hog farms, is now dominated by low cost large specialized farms (McBride and Key, 2003; Kaus, 2019). We believe that the recent reorganization of the hog industry could be, in part, due to spillover benefits (agglomeration economies) across states. In this study we examine the importance of agglomeration economies across states on the U.S. hog industry after controlling for factors that have been found to be important in shaping the reorganization of an animal production industry in prior studies. Our findings show that hog production in a state is positively affected by hog production in a nearby state, supporting the argument that agglomeration economies played an important role in the reorganization of the U.S. hog industry during the period of the study.

Hubbell and Welsh (1998) provided one of the earliest studies to formally address geographical concentration in the hog industry. Based on Theil’s entropy index, showed that hog production is becoming more geographically concentrated at the national level and within states. Herath, Weersink, and Carpentier (2005a), making use of the Gini coefficient to measure concentration, confirmed that hog production is becoming more concentrated within states. The changes associated with the hog industry becoming more geographically concentrated have led to a non-uniform interregional distribution of hogs (Hubbell and Welsh, 1998). The geographical concentration of hog farms can have several lasting effects. Specifically, the changes in the hog production industry could: affect communities with changes in hog production levels; affect local supply and demand for key inputs and output; alter the economic base of communities; change the utilization of industry-specific infrastructure and services; and concentrate nutrients from animal manure in fewer locations causing adverse environmental consequences (Roe, Irwin, and Sharp, 2002).

*Corresponding Author: Gibson Nene, Department of Economics, University of Minnesota Duluth.
Email: gnene@d.umn.edu, Tel: 218 726 7864
Co-Author Email: kschoengold2@unl.edu
The reorganization of an industry has been found to be affected by several factors including those internal to the firm; factors external to the firm but internal to the industry; and factors external to both the firm and the industry. Eberts and McMillen (1999) noted that firm-internal scale economies, industry internal scale economies (external economies to the firm), urbanization economies (external to both the firm and the industry), transportation costs, and environmental regulations are important factors affecting firm location. The U.S. hog industry is not unique, such factors were found to be important to the U.S. hog industry by Roe, Irwin, and Sharp (2002) for the year 1997 and between the years 1992 and 1997. Agglomeration economies are economies external to the farm but internal to the hog production sector. In this study, agglomeration economies are defined in the context of localization economies as all the benefits that accrue to hog production in one state as a result of hog production in a nearby state. Such spillovers are likely to arise because the presence of other hog feeding operations facilitates a local, industry-specific infrastructure of service individuals and information, which enhances the performance of each operation through reduced transaction costs and improved diffusion of production, financial and marketing information (Eberts and McMillen, 1999).

These factors are more than likely to have played an important role in the reorganization of U.S. hog production within and across states witnessed over the past two decades. Easy access to factors of production plays an important role on a farmer’s decision on whether or not to operate a hog operation facility in a particular state. The variation of input costs may lead to variations in input use through time. Because such fluctuations are inevitable, the cost of the factors of production will pose a great challenge to existing and potential hog producers. Because one of the main reasons why farmers engage in hog production is to make a profit, the attractiveness of the market plays an important role in determining whether to engage in hog production in a particular location or not. High market prices will more than likely attract more farmers to engage in hog production. Environmental regulations in the hog industry played important role in the reorganization of the U.S. hog industry since the 1990s. Several studies found evidence that environmental regulations have slowed down U.S. hog production (Metcalfe, 2001; Roe, Irwin and Sharp, 2002; Kuo, 2005; and Herath, Weersink, and Carpenter, 2005b; Azzam, Nene and Schoengold, 2015; Kaus, 2019) showed environmental regulations have an impact on the structure of the U.S. hog industry.

While the U.S. hog industry has received wide research attention in recent years, the importance of agglomeration economies on U.S. hog production has received limited attention with mixed results. To the best of our knowledge, only a single study by Roe, Irwin and Sharp (2002) addressed the importance of agglomeration economies on U.S. hog production. Accounting for agglomeration economies by making use of the spatial lag serves two distinct purposes; firstly, it corrects spatial autocorrelation through the dependent variable and secondly, this variable captures whether agglomeration economies are present or not. The study by Roe, Irwin and Sharp (2002) is based on a cross-section of counties for the years 1992 and 1997. Based on a cross-section of counties using 1997 data, the study finds that the presence of other swine farms has a positive effect on the inventory of hogs in a particular county. The study also found that hog production in one county to be negatively correlated with hog production in a nearby county.

The study by Roe, Irwin and Sharp (2002) is, however, limited in that it addresses the importance of agglomeration economies over a single year, 1997, and between two years, 1992 and 1997. Their findings that agglomeration economies existed in the hog industry in the year 1997 and such economies did not exist between 1992 and 1997, do not capture the importance of agglomeration economies over time. Sorting out the true effects as they exist and whether they have persisted or changed over time is important to fully understand such effects (Dean, Brown and Stango, 2000).

We believe that the importance of agglomeration economies and regulation in hog production is not likely to remain the same over time as such spillovers are likely to change due to several factors including changes in the stability of the economy stemming from recessions, depressions or booms. Economic hardships may disrupt or change the institutions central to the diffusion of production, marketing and financial information forcing farmers to go out of business. While the foregoing studies have been useful in the understanding of issues in U.S. hog production, to our knowledge, no study has investigated the importance of agglomeration economies on U.S. hog production by making use of panel data at the state level and the possibility of spatial dependence that stems from omitted variables that are related to each other over space. The current study fills this gap.

Our study will make use of panel data from top 22 U.S. major hog producing states to address the importance of agglomeration economies while controlling for the factors that have been found to be important in shaping the organization of an industry such as input availability, market attractiveness, and environmental regulations. Because most of the top hog producing states are in the Midwest, our study also examines the presence or absence of agglomeration economies in the Midwest states. To address this problem, we make use of a simple profit maximization model for a theoretical framework. The empirical analysis is based on three alternative models; ordinary least squares (OLS), two-stage least squares-spatial lag model (2SLS-SLM) and generalized method of moments-spatial autoregressive (GMM-SAR). Results from the three alternative models are then used to determine the model that best fits the data. While the
effects of the other factors can be examined using all three models, agglomeration economies can only be examined using the 2SLS-SLM model. Our study also tests for spatial autocorrelation through the error term using the GMM-SAR model. A review of spatial models is presented next.

REVIEW OF SPATIAL MODELS

Cross-sectional models that assume that the dependent variable corresponding to each cross-sectional unit depends, in part, on a weighted average of that dependent variable corresponding to neighboring cross-sectional units are called spatial autoregressive models (Anselin, 1988). This weighted average is called a spatial lag of the dependent variable. The spatially lagged dependent variable is usually correlated with the error term rendering the ordinary least squares estimator inconsistent, (Ord, 1975; Anselin, 1988; Anselin and Bera, 1998).

The standard spatial autoregressive model also known as the spatial lag model (SLM) is defined as follows:

\[
y = \lambda Wy + X\beta + \mu
\]

\[
\mu \sim N(0, \sigma^2 I_n) \quad \text{,} \quad |\lambda| < 1 \quad \text{Equation (1)},
\]

where \( Y \) is the n x 1 vector of observations on the dependent variable, \( X \) is the n x k matrix of observations on k exogenous variables, \( W \) is an n x n spatial weighting matrix of known constants, \( \beta \) is the k x 1 vector of regression parameters, \( \lambda \) is a scalar autoregressive parameter, and \( \mu \) is the n x 1 vector of disturbances.

If the researcher believes that the spatial dependence stems from omitted variables that are related to each other over space, the spatial autoregressive (SAR) model, is employed. The SAR model is specified as follows:

\[
y = X\beta + \mu
\]

\[
\mu = \rho W\mu + \varepsilon \quad \text{,} \quad |\rho| < 1 \quad \text{Equation (2)},
\]

\[
\varepsilon \sim N(0, \sigma^2 I_n)
\]

where \( \mu \) is the n x 1 vector of disturbances, \( \rho \) is the autoregressive parameter and \( \varepsilon \) is the n x 1 vector of innovations. The rest of the parameters in the SAR model are as defined in the model in (1). Kelijian and Prucha (1999) proposed a “generalized” moments (GMM) approach for the estimation of the spatial autoregressive parameter, \( \rho \), which was found to be as efficient as the standard maximum likelihood estimator which was the popular method used prior to the GMM approach. The (quasi) maximum likelihood and the moments estimator of \( \rho \) was proposed by Ord (1975), and has been popular since.

The models in equations (1) and (2) if specified in a single model, the combined model is called the spatial autoregressive model with autoregressive disturbances (Kellijan and Prucha, 1998). This model is specified as follows:

\[
y = \lambda Wy + X\beta + \mu
\]

\[
\mu = \rho M\mu + \varepsilon
\]

\[
\varepsilon \sim N(0, \sigma^2 I_n)
\]

where \( y \) is the n x 1 vector of observations on the dependent variable, \( X \) is the n x k matrix of observations on k exogenous variables, \( W \) and \( M \) are n x n spatial weighting matrices of known constants, \( \beta \) is the k x 1 vector of regression parameters, \( \lambda \) and \( \rho \) are scalar autoregressive parameters, \( M \) is the n x 1 vector of disturbances, and \( \varepsilon \) is a nx1 vector of innovations. If we let \( \rho = 0 \), the model in equation (3) reduces to the model represented in equation (1), and when \( \lambda = 0 \) the model collapses into the model in equation (2). When \( \lambda = \rho = 0 \), then there is no presence of spatial correlation in the spatial autoregressive model with autoregressive disturbances.

A generalized spatial two-stage least squares (GS2SLS) procedure for estimating the spatial autoregressive model with autoregressive disturbances was developed by Kelijian and Prucha (1998). The GS2SLS model allows for the possibility that the spatial weight matrix associated with the error term and that associated with the dependent variable is the same, that is, \( M = W \). The GS2SLS approach is a three-step procedure. In the first step the regression model in equation (3) is estimated by two-stage least squares (2SLS) using instruments.

The instrument matrices used are a subset of the linearly independent columns of \((X, WX, W^2X, ..., MX, MWX, MW^2X)\), where the subset contains the linearly independent columns of \((XM, X)\). The autoregressive parameter, \( \rho \), is estimated by generalized method of moments (GMM) procedure suggested by Kelijian and Prucha (1998) in the second stage. The GMM estimation procedure yields a consistent estimator of \( \rho \), whether or not the weight matrices for the dependent variable and the error term are equal (Kelijian and Prucha, 1998). The third step, involves a 2SLS re-estimation of the model in (3) after transforming the model using a Cochrane-Orcutt type transformation to account for spatial correlation as follows:

Hog Production and Agglomeration Economies: The Case of U.S. State-Level Hog Production
First, equation (3) is rewritten as follows:
\[ y = Z\delta + \mu \]  
Equation (4),
where \( Z = (X, Wy) \) and \( \delta = (\beta', \lambda') \). The Cochrane-Orcutt type transformation to this model yields:
\[ y^* = Z^*\delta + \epsilon \]  
Equation (5),
where \( y^* = y - \rho Wy \) and \( Z^* = Z - \rho WZ \). In sum, the foregoing spatial models presented in equations (1), (2), and (3) are the popularly used models in spatial econometrics today. The outline of our theoretical model is presented next.

THEORETICAL MODEL

For our theoretical model, we present a simple profit maximization model applied to hog production. Hog production is assumed to utilize land, feed, labor, transportation, and the environment as inputs. The production function associated with hog production is therefore given by: 
\[ y = f(C, L, N, T, E), \] where \( y, C, L, N, T \) and \( E \) represent the quantity of output, the quantity of corn feed, the quantity of land, the transportation input and the quantity of the environmental input, respectively. The environment is treated as an input in this model because hog producers incur a cost to utilize the environment in disposing of hog waste including manure. The production of hogs is costly, and the total cost of production is given by
\[ V = w_c C + w_L L + w_N N + w_T T + w_E E \]
where, \( V, w_c, w_L, w_N, w_T \) and \( w_E \) represent total cost, hog price, price of feed, price of labor, price of land, transportation cost and environmental cost, respectively. The firm’s profit maximization problem in each state is given by:
\[ \max_{C, L, N, T, E} \pi = py - w_c C - w_L L - w_N N - w_T T - w_E E \]  
Equation (6)

Solving the problem above yields the following first order conditions:
\[ pf_c - v_{w_c} = 0 \]
\[ pf_L - v_{w_L} = 0 \]
\[ pf_N - v_{w_N} = 0 \]
\[ pf_T - v_{w_T} = 0 \]
\[ pf_E - v_{w_E} = 0 \]  
Equation (7)

By plugging in the input demand functions back into the profit function yields the indirect profit function given by:
\[ \pi^* = \pi(p, w_c, w_L, w_N, w_T, w_E) \]  
Equation (9)

The usual properties of the profit function are assumed to hold. The profit function is assumed to be homogenous of degree one in output and input prices. The partial derivative of the indirect profit function with respect to the output price, \( \frac{\partial \pi^*}{\partial p} = y > 0 \). This inequality satisfies the property that the profit function is non-decreasing in output price. The partial derivatives of the indirect profit function with respect to input prices are negative by the envelope theorem, that is, \( \frac{\partial \pi^*}{\partial w_i} = -i^* < 0 \), for \( i = c, L, N, T \) and \( E \).

By Hotelling’s Lemma, we can derive the optimal hog supply function for hogs from the indirect profit function given in equation (9), i.e. \( \frac{\partial \pi^*}{\partial p} = y^* \). The hog supply function is a function of the output price and all input prices. We can write the hog supply function as follows:
\[ y^* = y(p, w_c, w_L, w_N, w_T, w_E) \]  
Equation (10)

The effect of environmental costs on hog production will be captured by how environmental stringency affects hog supply. Given that we model environmental regulations as an input in the production of hogs, it follows that the comparative statics results of this effect on profit and the optimal hog supply are given by \( \frac{\partial \pi^*}{\partial w_E} < 0 \), and \( \frac{\partial y^*}{\partial w_E} < 0 \), respectively. This stems from one of the properties of the profit and supply functions that they are non-increasing in input prices. In Section 4 we present an empirical model that estimates the optimal supply function. Given the above comparative static results, theory predicts that the effect of environmental regulations on hog production is negative.
EMPIRICAL MODEL

In this study we examine the importance of agglomeration economies on hog production for a panel consisting of top 22 hog producing states in the U.S. We use the Gauss software to estimate a menu of models based on pooled OLS, the 2SLS-SLM presented in equation (1) using the two-stage least squares procedure developed by Kelijian and Prucha (1998), and the GMM-SAR model presented in equation (2) using the generalized method of moments (GMM) developed by Kelijian and Prucha (1999). We use the same empirical model specification to estimate results for the top 22 hog producing states to those based on Midwest hog producing states.

If the 2SLS-SLM and the GMM-SAR models show that spatial autocorrelation through both the dependent variable and the error term exist, we will employ the GS2SLS model which captures the two spatial measures in a single model. If results show the existence of spatial autocorrelation exists only, the best model for the data will be the 2SLS-SLM model. On the other hand if spatial autocorrelation exists only through the error term, then the best model for the data will be the GMM-SAR model. The OLS model estimates will be reliable in the event that spatial autocorrelation is non-existent both through the dependent variable and the error term. Our strategy is to first check for the importance of all possible spatial autocorrelation sources to come up with the model that best explains the data rather than impose a model apriori.

We rewrite the components of equations (1), (2) and (3) which are specified for cross-section data so that they suit the panel data model. We rewrite these components as follows: the dependent variable, $y$, is the $nt \times 1$ vector of observations on hog output; $X$ is the $nt \times k$ matrix of observations on k exogenous variables; $W$ is an $nt \times nt$ spatial weighting matrix of known constants, $\beta$ is the $k \times 1$ vector of regression parameters, $\lambda$ and $\rho$ are scalar autoregressive parameters, $\mu$ is the $nt \times 1$ vector of disturbances, and $\iota$ is the $nt \times 1$ vector of innovations. For the spatial weighting matrix, we use a standardized first-order spatial contiguity matrix. We assign a “1” for states that share a common border and a “0” for states that do not share a common border. Notice that a state cannot be its own neighbor, henceforth; the elements on the main diagonal of the first-order contiguity matrix are all set to zero. The spatial weight matrix is standardized by normalizing so that row sums add to unity (LeSage, 1997). Several studies including Pan and LeSage (1995) and LeSage (1997), have used the first-order contiguity matrix in spatial econometrics.

The menu of spatial models considered in this study is given below:

\[ y = \lambda Wy + X\beta + \mu \]
\[ \mu \sim N(0, \sigma^2 I_{nt}) \]
\[ |\lambda| < 1 \]  \hspace{1cm} \text{Equation (1')},

\[ y = X\beta + \mu \]
\[ \mu = \rho W\mu + \epsilon \]
\[ |\rho| < 1 \]  \hspace{1cm} \text{Equation (2')},

\[ \epsilon \sim N(0, \sigma^2 I_{nt}) \]

\[ y^* = Z^* \delta + \epsilon \]  \hspace{1cm} \text{Eq} \hspace{1cm} \text{Equation (5')},

where models in equations (1’), (2’) and (5’) are panel data versions for models in equations (1), (2) and (5), respectively.

The dependent and independent variables used in this study are defined next. The dependent variable, $y$, is the state level percentage share of U.S. total hog production. The state level percentage share of U.S. total hog production is used to capture state-level hog supply as in Metcalfe (2001). Roe, Irwin and Sharpe used the logarithm of U.S. hog production as the dependent variable which is not far from what we use here.

The specific independent variables contained in the $X$ matrix include: hog output price; the cost of factors of production variables, corn price, farm labor, transportation cost, land price; and the environmental input cost (index). The hog output price is endogenous in our model as it is associated with the demand side of the industry. We make use of the predicted values for hog output price to take care of the endogeneity problem as in Metcalfe (2001). A positive sign is expected since the higher the price of hogs, the more hog suppliers are willing to supply hogs. Corn price reflects the cost of the corn input in hog production which constitutes a greater percentage of hog feed. Several studies used corn price as an input to study the effect of environmental regulations on hog production (Metcalfe, 2001; Roe, Irwin and Sharp, 2002). We expect a negative sign on this variable since the higher the corn feed input price the lower the number of hogs hog producers are willing to supply.

To capture the cost of farm labor in hog production, we use farm labor wages. Specifically we use the average hourly state-level wages as a proxy for farm labor wages. Metcalfe (2001), and Herath, Weersink, and Carpentier (2005b) also used farm labor wages to capture the cost of farm labor input. A negative sign on this variable is expected, that is, the higher the farm labor cost, the lower the number of hogs hog producers are willing to supply. Land price captures the cost of land input in hog production. We use farm land price to capture land input cost in the same manner as in Metcalfe (2001), and Herath, Weersink, and Carpentier (2005b). We expect the coefficient on the land price variable to be negative. A negative sign shows that hog farmers supply fewer hogs as the price of land input goes up. Transportation cost
reflects the cost of transport input in hog production. We capture transportation cost using the price of unleaded gas. Energy price is a popular proxy for transportation cost in studies in livestock industry studies as well as meatpacking industry studies, e.g. Azzam (1997), Metcalfe (2001) and Herath, Weersink, and Carpenter (2005b) used the price of gas as a proxy for transportation cost in hog related studies. A negative sign is expected for this variable since it is a cost of an input in the hog production process.

For the environmental index variable, we make use of a time series of qualitative environmental stringency indices constructed in the same manner as in Metcalfe (2000). Qualitative environmental stringency indices have been widely used in hog industry studies (Metcalfe, 2001; Roe, Irwin, and Sharp, 2002; Herath, Weersink, and Carpenter, 2005b; Azzam, Nene and Schoengold, 2015). Specifically, we construct the indices for the years 2003 to 2006, and we combine these with indices constructed in prior studies (Metcalfe, 2000; and Herath, Weersink, and Carpenter, 2005b), to complete the 1994-2006 time series for each of the top 22 hog producing states.

Because stringency indices used in prior studies are based on different measures and judgments, we make use of the method employed by Herath, Weersink, and Carpenter (2005b), to go around the problem of different stringency indices that exist in the literature. The methodology uses the ratio of the state value divided by the mean of the state’s regulation stringency for the period in the sample. If the value of this ratio is greater than 1, equals 1 and is less than 1, that state has above average, average and below average stringency level, respectively. This is one way to make the different stringency measures comparable. Environmental stringency is endogenous in our model because regulations may increase in states that are experiencing increasing hog production while states with low environmental regulations may realize increased hog production (Metcalfe, 2001). Predicted values of the environmental stringency measure are used in order to take care of this endogeneity. A negative sign is expected on this variable as suggested by theory, \(\frac{\partial y}{\partial w} < 0\), as it enters the hog production as an input. Specifically, in this model the indices reflect an environmental regulation compliance cost.

Spatial lag (\(\lambda\)) is used as a proxy for agglomeration economies. Specifically, this variable captures spatial autocorrelation thorough the dependent variable which we call agglomeration economies in this study. We define this variable in the same manner as in Roe, Irwin and Sharp (2002). A positive (negative) and statistically significant sign on the spatial lag variable suggests the presence (absence) of agglomeration economies in the hog industry. A positive or a negative sign is expected. The scalar autoregressive parameter, \(\rho\), is used to capture spatial dependence through the error term. While this is not a parameter of interest, if ignored when spatial autocorrelation exists through the error term, our results will be inconsistent. Data sources and units of study are presented next.

**DATA**

The study uses data for 22 major U.S. hog producing states for the years 1994 through 2006. The period of the study is limited by the availability of environmental regulation stringency index data which is only available from 1994 to 2006. We also estimate regressions for the 22 top U.S. hog producing states and for the Midwest hog producing states. We separate these for hog producing states from the rest of the 22 states for several reasons. First, spatial dependence in the Midwest is likely to be different from the spatial dependence we observe when all the 22 states are considered. Secondly, the Midwest states account for most of the corn production in the U.S. which is a major input in hog production, and thirdly, the Midwest states include eight of the top 10 U.S. hog producing states; Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, and Ohio. The 22 top U.S. hog producing states account for about 85% of total U.S. hog production. The 22 hog producing states and the Midwest hog producing states are shown in Table 1 below.

<table>
<thead>
<tr>
<th>Unit of study</th>
<th>States</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 major producing states</td>
<td>Arkansas, Colorado, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Carolina, Ohio, Oklahoma, Pennsylvania, South Dakota, Tennessee, Texas, Utah, Virginia, and Wisconsin</td>
</tr>
<tr>
<td>Midwest states</td>
<td>Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, South Dakota, and Wisconsin</td>
</tr>
<tr>
<td>Arkansas, Colorado, Georgia, Kentucky, North Carolina, Oklahoma, Tennessee, Texas, Utah, and Virginia</td>
<td></td>
</tr>
</tbody>
</table>

The data sources and description, and the descriptive statistics are given in Tables 2, and 3, respectively.
Table 2. Variable Definition and Data Sources

<table>
<thead>
<tr>
<th>Definition of variables</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hog farms total inventory</td>
<td>USDA-NASS</td>
</tr>
<tr>
<td>Land value: Dollars/acre</td>
<td>NASS Agricultural Prices Summaries- USDA</td>
</tr>
<tr>
<td>Hog price: Dollars/Cwt</td>
<td>NASS Agricultural Prices Summaries- USDA</td>
</tr>
<tr>
<td>Corn price: Dollars/Bushel</td>
<td>NASS Agricultural Prices Summaries- USDA</td>
</tr>
<tr>
<td>Transportation cost: cents per gallon:</td>
<td>Energy Information Administration</td>
</tr>
<tr>
<td>Farm labor: Dollars/Hour</td>
<td>Farm labor-NASS-USD</td>
</tr>
<tr>
<td>Cattle all beef price: Dollars/Cwt</td>
<td>NASS Agricultural Prices Summaries- USDA</td>
</tr>
<tr>
<td>Income: personal per capita income/Dollars</td>
<td>Almanac of the 50 States Information Publications</td>
</tr>
<tr>
<td>Population density(pd): persons per square mile</td>
<td>Almanac of the 50 States Information Publications</td>
</tr>
</tbody>
</table>

Note for Table 2. Cwt refers to 100 weight

Table 3. Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>StdDev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output%</td>
<td>4.40</td>
<td>5.84</td>
<td>0.07</td>
<td>27.68</td>
</tr>
<tr>
<td>Hog price</td>
<td>42.13</td>
<td>8.72</td>
<td>27.47</td>
<td>83.51</td>
</tr>
<tr>
<td>Corn price</td>
<td>2.43</td>
<td>0.56</td>
<td>1.5</td>
<td>4.38</td>
</tr>
<tr>
<td>Transport</td>
<td>101.27</td>
<td>31.99</td>
<td>58.53</td>
<td>182.71</td>
</tr>
<tr>
<td>Labor</td>
<td>7.88</td>
<td>0.80</td>
<td>6.16</td>
<td>9.96</td>
</tr>
<tr>
<td>Land</td>
<td>1585.57</td>
<td>817.92</td>
<td>332.32</td>
<td>4185.42</td>
</tr>
<tr>
<td>Population density</td>
<td>110.56</td>
<td>78.05</td>
<td>9.5</td>
<td>280.3</td>
</tr>
<tr>
<td>Environmental stringency</td>
<td>1.00</td>
<td>0.52</td>
<td>0.2</td>
<td>2.38</td>
</tr>
<tr>
<td>Cattle price</td>
<td>65.49</td>
<td>10.83</td>
<td>32.76</td>
<td>96.99</td>
</tr>
<tr>
<td>Per capita income</td>
<td>27131.3</td>
<td>2981.53</td>
<td>20329.36</td>
<td>35334.31</td>
</tr>
</tbody>
</table>

EMPIRICAL RESULTS AND DISCUSSION

We examine the importance of agglomeration economies on the hog industry on all the top 22 hog producing states and on the Midwest hog producing states for robustness check. Results for the top 22 hog producing states, and Midwest hog producing states, are given in Tables 4, and 5, respectively. Table 4 below provides results for the 22 major hog producing states based on three different models, namely; 2SLS-SLM, GMM-SAR and OLS.

Table 4. Results for the 22 major hog producing states

<table>
<thead>
<tr>
<th>Variable</th>
<th>2 SLS-SLM</th>
<th>GMM-SAR</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>T-ratio</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Hog price</td>
<td>20.04</td>
<td>3.86***</td>
<td>25.19</td>
</tr>
<tr>
<td>Corn price</td>
<td>-33.81</td>
<td>-6.68***</td>
<td>-39.86</td>
</tr>
<tr>
<td>Transport</td>
<td>-1.66</td>
<td>-0.64</td>
<td>-3.20</td>
</tr>
<tr>
<td>Labor</td>
<td>31.47</td>
<td>4.08***</td>
<td>47.99</td>
</tr>
<tr>
<td>Land</td>
<td>7.23</td>
<td>6.43***</td>
<td>8.18</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.57</td>
<td>10.41***</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.37</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

Note for Table 4: The dependent variable is the state-level percentage share of total U.S. hog production. P values are indicated as ***0.01, **0.05 and *0.10.

The 2SLS-SLM results in Table 4 show that the coefficient of $\lambda$ is positive and statistically significant while the coefficient of $\rho$ is positive and statistically insignificant. Based on the GMM-SAR model results, we reject the null hypothesis that spatial autocorrelation through the error term exists at the state-level. Since there is evidence of positive spatial autocorrelation, the OLS model results are unbiased but inefficient (Anselin, 1988). We therefore conclude that the model that best explains our data is the 2SLS-SLM.

Based on the 2SLS-SLM results, the coefficients on hog price, corn price, and transport have the expected signs. The coefficient on hog price suggests that hog production increased with an increase in the price of hogs. The
coefficients on corn price and transport suggest that high prices of these factors of production serve to deter hog production in the 22 major hog producing states. The coefficients on labor and land have unexpected positive and statistically significant signs. The coefficient on labor suggests that the presence of a large pool of labor or relatively low farm labor wages serve to attract hog production. The positive coefficient on land suggests that land values are favorable for hog production probably due to the availability of vast farming land. The positive and statistically significant coefficient on $\lambda$ suggests that production in a state is positively affected by hog production in a nearby state, confirming the existence of agglomeration economies in the U.S. hog industry. The coefficient on index is negative and statistically significant. This result suggests that environmental regulation has a negative effect on hog production. Results for the Midwest hog producing states in Table 5 below also show that the best model for the data is the 2SLS-SLM.

Table 5. Midwest hog producing states

<table>
<thead>
<tr>
<th>Variable</th>
<th>2 SLS Spatial Lag</th>
<th>GMM-SAR</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>T-ratio</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Constant</td>
<td>-112.83</td>
<td>-6.22***</td>
<td>-104.35</td>
</tr>
<tr>
<td>Index</td>
<td>-116.10</td>
<td>-9.16***</td>
<td>-116.05</td>
</tr>
<tr>
<td>Hog price</td>
<td>21.95</td>
<td>2.48***</td>
<td>7.69</td>
</tr>
<tr>
<td>Corn price</td>
<td>-62.31</td>
<td>-6.99***</td>
<td>-63.27</td>
</tr>
<tr>
<td>Transport</td>
<td>2.10</td>
<td>0.55</td>
<td>6.27</td>
</tr>
<tr>
<td>Labor</td>
<td>62.38</td>
<td>5.33***</td>
<td>71.34</td>
</tr>
<tr>
<td>Land</td>
<td>10.96</td>
<td>6.51***</td>
<td>12.86</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.35</td>
<td>5.03***</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note for Table 5: The dependent variable is the state-level percentage share of total U.S. hog production. P values are indicated as ***0.01, **0.05 and *0.10.

Based on 2SLS-SLM results in Table 5, hog production in a state is positively affected by hog production in a nearby state and environmental regulation has a negative effect on hog production. This result suggests that agglomeration economies are important for the Midwest hog producing states. The coefficients on hog price, and corn price have the expected signs, suggesting that hog price attracts hog production while the corn input cost deters hog production in the Midwest. The coefficient on transport is positive and statistically insignificant. This result suggests that transport cost has no effect on hog production in the Midwest. The coefficients on land and labor have positive and statistically significant signs suggesting that the costs of these factors of production are favorable for hog production in the Midwest. The coefficient on index is negative and statistically significant. This result suggests that environmental regulations have a negative effect on hog production. We established that the model that describes our data better for the 22 states, and Midwest states, is the 2SLS-SLM. We therefore document only the 2SLS-SLM results already reported in Tables 4 and 5. These results are given in Table 6 below.

Table 6. 22 States and Midwest 2SLS-SLM Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>All states</th>
<th>Midwest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>T-ratio</td>
</tr>
<tr>
<td>Constant</td>
<td>-73.50</td>
<td>-6.44***</td>
</tr>
<tr>
<td>Index</td>
<td>-58.50</td>
<td>-7.75***</td>
</tr>
<tr>
<td>Hog price</td>
<td>20.04</td>
<td>3.86***</td>
</tr>
<tr>
<td>Corn price</td>
<td>-33.81</td>
<td>-6.68***</td>
</tr>
<tr>
<td>Transport</td>
<td>-1.66</td>
<td>-0.64</td>
</tr>
<tr>
<td>Labor</td>
<td>31.47</td>
<td>4.08***</td>
</tr>
<tr>
<td>Land</td>
<td>7.23</td>
<td>6.43***</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.57</td>
<td>10.41***</td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note for Table 8: The dependent variable is the state-level percentage share of total U.S. hog production. P values are indicated as ***0.01, **0.05 and *0.10.

Results in Table 6 show that agglomeration economies are stronger for the 22 major hog producing states when compared to Midwest hog producing states. This suggests that hog production in a state is more positively related to that of a neighboring state, when the 22 major producing states are considered. This difference could be due to the fact that the non-Midwestern states do benefit a lot from being close to Midwestern states as they do not have the large supplies of corn that Midwestern states have and other well-established infrastructure tailored for hog production.
production. In general, results confirm that agglomeration economies play a major role in U.S. hog production regardless of the region being considered. The presence of agglomeration economies across states shows that hog production across states benefits from hog industry-specific infrastructure of service individuals and information. This in turn enhances the performance of each hog production operation through reduced transaction costs and improved diffusion of production, financial and marketing information. Results on the effect of the cost of factors of production on the U.S. hog production industry suggest that: transport cost has no effect on hog production; land value and labor cost have a positive effect on hog production; and the cost of corn input has a negative effect on hog production for the 22 major hog producing states, and the Midwest hog producing states.

CONCLUSIONS

This study empirically examined the importance agglomeration economies on the U.S. hog industry for the period 1994-2006 based on a simple profit maximization theoretical model. The analysis used data from the top 22 U.S. hog producing states and the analysis was replicated for the Midwest hog producing states for a robustness check. We estimated three empirical models, Pooled OLS, 2SLS-SLM, and GMM-SAR in this study. The GMM-SAR model results show that the error space parameter is statistically insignificant in all the GMM-SAR based models. Based on the GMM-SAR model results, we reject the null hypothesis that spatial autocorrelation through the error term exists at the state-level. The 2SLS-SLM model results show that the coefficient on the agglomeration economies variable is highly statistically significant in all model specifications. The presence of agglomeration economies in the 2SLS-SLM rules out the OLS results, Anselin (1988). We therefore conclude that of the three models, the model that best explains our data is the 2SLS-SLM.

The 2SLS-SLM results suggest that hog production in a state is positively affected by hog production in a nearby state over time, confirming the presence of agglomeration economies in the U.S. hog industry. This result holds for both the 22 major hog producing states and the Midwest hog producing states. This is a similar result to the county-level cross-section result based on 1997 data found by Roe, Irwin, and Sharp (2002). However, this result contrasts with the result by Roe, Irwin, and Sharp (2002) on changes in regulation between the years 1992-1997. Indeed, their result may have failed to capture the changes in regulation over time. Our results complement results from existing studies by showing that agglomeration economies are also important over time in the U.S. hog industry.

Policy makers in one state should consider working with policy makers neighboring states in order to come up with policies which are mutually beneficial for hog production in order to make it easy for hog producers in different states to take advantage of agglomeration economies through ease access to inputs, information sharing, diffusion of production, reduced transaction costs, and knowledge spillovers, among others. An area for further investigation would be trying to answer the question: Is environmental regulation in a state influenced by regulation in nearby states? States may consider regulation in nearby state as a way to reduce the time cost of coming up with new regulations or to make their regulation more stringent than those of nearby states for political reasons.

REFERENCES


**Accepted 16 December 2019**


**Copyright:** © 2019: Nene and Schoengold. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.