Title: Integration of Regional Land Markets in Texas

Name of authors: Tian Su

Senarath Dharmasena

David Leatham

Charles E. Gililland

Affiliation: Department of Agricultural Economics, Texas A&M University
Real Estate Center, Texas A&M University

Selected Paper Prepared for Presentation at the Southern Agricultural Economics Association (SAEA) Annual Meeting, Louisville, Kentucky, February 1-4, 2020.

Abstract

Texas contains diverse landscapes varying from deserts to piney woods and from sea-level plains to alpine vistas. These diversities create varying conditions in land markets which yield markets providing more services than simply food or natural resource production. Farmland, besides providing fertile fields for growing crops, also acts as a home for many farmers and land investors. The natural amenities such as lakes and forests have been widely recognized. Practicing appraisers have classified Texas rural land markets into distinct regions based on market trends and geographical features. Given the plethora of factors determining the value of land markets in Texas, this study uses cutting-edge machine learning and graphical methods and posted land prices to uncover interaction patterns of land markets of different regions of Texas. This information will be useful for various stakeholders in regional land markets to strategically position land transactions based on integration patterns of regional land markets. Preliminary study of data reveals that land price in Region 6 (South TX) is determined by those of Regions 7 (Hill Country), Region 1 (Panhandle/South Plains), Region 2 (Far West Texas) and Region 4 (Northeast TX). Land price of Region 4 also determines that of Region 7 and Region 2 via Region 7.

1. Introduction

Rural land[1] markets and the factors that affect the value of such markets have attracted the attention of both academics and practitioners. The diversity of physical characteristics of land and its uses place land in a vital position regarding social, economic and environmental issues. Among the 2,200 million acres of land in the United States, 97% is categorized as rural land (USDA ERS-MLU). Moreover, the most common use of rural land is farmland, which amounts to over 40% of the total rural land use (USDA ERS-MLU). In general, rural land uses and values indicate the distribution of benefits from relevant economic activity and provide information to guide business decisions.

Texas rural land markets have special features that make it unique from other rural land markets in the U.S. First, Texas is the second largest state with the second largest population in the U.S.

[1] Rural land in this work follows the Census Bureau urban and rural area definition. Related details can be found at the census website:https://www2.census.gov/geo/pdfs/reference/GARM/Ch12GARM.pdf.

Based on a report[2] conducted by Texas Land Trends, 83% of the state's land area is rural land. There is a massive land market in Texas, where a huge amount of transactions are made annually (Gilliland, 2012). Second, due to its size and geologic features, Texas contains diverse landscapes. Those variations of landscapes create widely divergent conditions that affect the ownership and marketing of land (Gilliland et al. 2012). Third, Texas is rich in crude oil and natural gas. Texas has 39% of the proved crude oil reserves and 26% of the proved gas reserves in the U.S[3]. Also, the crude oil and natural gas production are widely distributed in Texas. Of the 254 counties in Texas, 223 have oil and/or natural gas production.[4] The land market will inevitably be affected by the natural resources beneath the land. Fourth, Texas has the most farms and the highest farm acreage in the U.S in 2015-2016(USDA NASS). The state ranked as the first for revenue generated from total livestock and livestock products in 2017 (USDA ERS-FIWS). It ranked third for total agricultural revenue in 2017 (USDA ERS-FIWS). These extensive agricultural uses of rural land significantly impact land market. Besides providing substantial economic, environmental, and recreational resources that benefit many Texans, the abundant number of rural land sales act as good indicators for the overall rural land market. Therefore, a study of the rural land market is relevant.

Only a few studies emphasize land price or valuation specific to Texas rural land markets. Among those, Faubion (1971) found some promising economic factors that could predict future land value. Although his results were favorable, questions of multicollinearity exist. Moreover, because he only applied data from El Paso County, the study is not necessarily applicable to Texas as a whole. To the author's best of knowledge, Pope (1985) is the only comprehensive study on the Texas rural land markets. Pope (1985) estimated several factors affecting rural land values. However, his conclusion might be less applicable due to limited transaction data. Although there are a few empirical studies concerning rural land markets in other states, they are limited by the size of their available datasets. Examples include Kennedy et al. (1997) and Jones et al. (2006). Therefore,

[2] *Status Update and Trends of Texas Rural Working Lands*(2017) found at https://medium.com/texas-land-trends/status-update-and-trends-of-texas-rural-working-lands-2ba34c3d4963.
[3] Data available at: https://www.eia.gov/dnav/pet/pet_crd_pres_dcu_NUS_a.htm; https://www.eia.gov/dnav/ng/ng_enr_sum_dcu_STX_a.htm.
[4] Data available at: https://www.energyindepth.org/wp-content/uploads/2010/07/TX-Oil-and-Nat-Gas.pdf

specialized research investigating Texas rural land market proves to bridge a gap in existing literature.

As indicated by economists from Real Estate Center at Texas A&M, Texas rural land markets are classified into distinct regions based on market trends and geographical features that practicing appraisers have cited as market influences. Figure 1 provides a detail map of seven land regions in Texas. The boom and bust of regional markets and price levels are primarily determined by local economic fundamentals as well as the national, macro economy. Changes in markets could also result from positive or negative trends in other areas of the economy. A comprehensive market integration study is conducted in this paper to understand the interdependencies across Texas land markets. The conclusions of this study allow lenders and policymakers to anticipate changes in trends in one region of Texas resulting from a change observed in other regions and thus allows for more sound decision making.

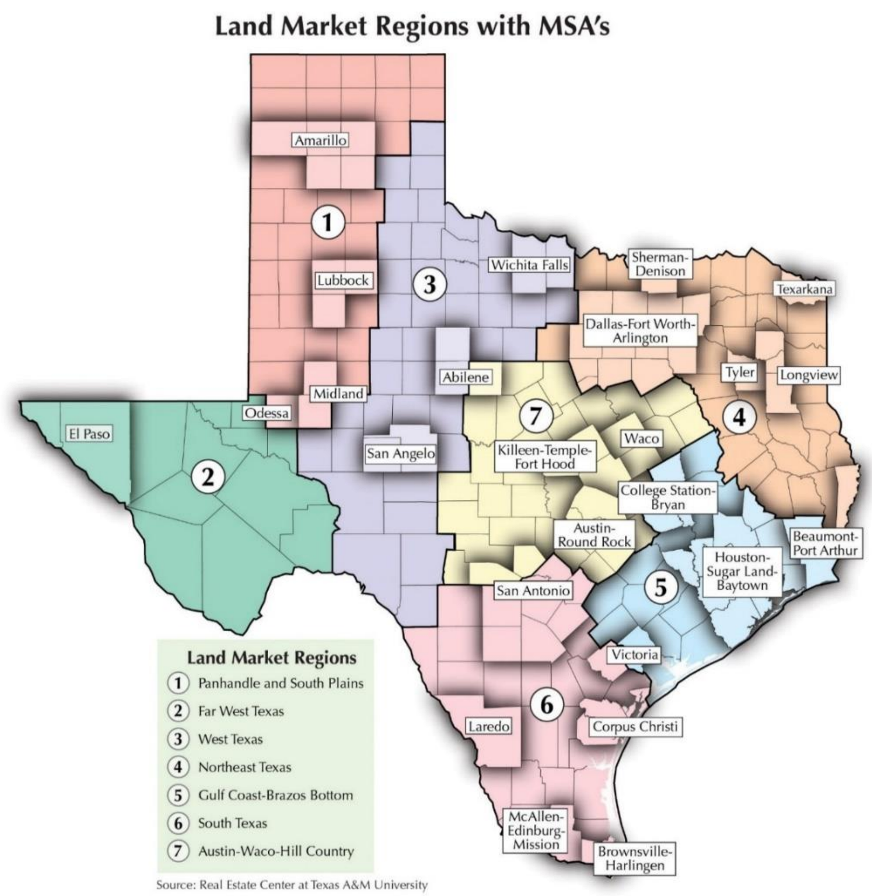

Figure1. Texas Land Market Regions

The paper is organized as follows. Section 2 presents a brief literature review. Section 3 discusses the analytical framework and data, followed by section 4 dealing with empirical results. Section 5 is the conclusion and implications for future research.

2. Brief Literature Review

According to the best of knowledge of the authors, a comprehensive study investigating the interrelationships between the different rural land markets in Texas is not available in the extant literature. There exists a series of related regional housing market analyses which are presented in this section. It is reasonable to expect a similar phenomenon from these studies in rural land markets. Studies of interrelationships between prices in different markets are also reviewed because methodologies applied in this literature shed light on the analytical framework used.

Empirical Explorations of Market Integration

Correlations between house prices at different locations have attracted extensive concerns (Zhang, L. et al.,2017). Many empirical studies use time series methods to analyze house price diffusion. MacDonald and Taylor (1993) and Alexander and Barrow (1994) find that long-run co-integration relationships exist between regions in Great Britain. Munro and Tu (1996) proposed a Johansen co-integration technique to model the dynamics of both national and regional house prices. Considering spatial interaction, house prices in one region are included in the price equations of other regions. Pollakowski and Ray (1997) use the Vector Auto Regression (VAR) model and Granger causality to show that house price diffusion exists between the five largest primary metropolitan statistical areas in the Greater New York area. In addition, a strong lead-lag correlation between house price changes among Taiwan's largest urban areas is observed by Lee, and Lin (2014). Similar studies have been applied to Ireland, Australia, New Zealand, and South Africa, where co-integration and the Granger Causality test are used to study the long-run equilibrium relationships among regional housing markets. (Lee. and Chien, 2011; Luo, Liu, & Picken, 2007; Shi et al., 2009; Stevenson, 2004).

While the aforementioned studies explain the ripple effect by space dependence, some scholars argue that differences in structural variables also play a role in the ripple effect. Meen (1999) propose a coefficient heterogeneity model, indicating that structural or behavioral differences in the regional markets can lead to the ripple effect. Holmes and Grimes (2008) and Tsai (2014) found that if the ripple effect exists, there would be a long-run stable relationship between regional house prices, which is consistent with Meen’s (1999) results. Considering coefficient heterogeneity mentioned in Meen (1999), a regional house price model with determinants including national house price, interest rates and other economic variables in each region is proposed by Zhang, Gerlowski, and Ford (2014). The result shows that house prices in one region are influenced by economic variables in other regions.

Previous Methodological Applications

Besides time series models mentioned in housing market studies, a data driven method that focuses on inferring causality is beginning to emerge documenting the market interaction studies. The suggested method of directed graphs relies on algorithms that generate causal relationships among observational data. Directed acyclic graphs (DAGs) are the graphical forms used to present the causal structures derived from these algorithms. Machine learning techniques play a vital part in constructing the structure of these graphs. These techniques allow one to make inferences when the interactions among variable are too complex (Senia et al., 2018; Yu et al., 2019).

Although few studies on rural land market exists that apply DAG, the graphical analysis has proved to be reliable in many other economic areas. Bessler and Akleman (1998) employ the DAG theory to modeling pork and beef price at both farm and retail levels. Their study helps to providing standard practice of DAG by assuming universal information set is causally sufficient. Similar work has also been completed in other spatially separated markets. For example, Fang and Bessler (2017) contribute to the Asian stock market by finding the lead lag relationships during the market collapse in 2015. Rettenmaier and Wang (2012) and Dhamarsena et al. (2016) are good applications of DAG by inferring possible causality in complex environments. By implementing DAGs to examine causal factors of health outcomes, Rettenmaier and Wang (2012)'s work successfully distinguish causality from correlation. Dhamarsena et al. (2016) realize the complexity of the food environment. They select several factors discussed in previous literature and generate a holistic picture of the "food environment" by applying DAGs.

Overall, previous studies indicate that price interactions exist in many markets around the world. It is reasonable to expect a similar phenomenon in rural land markets. Therefore, this paper serves to fill the gap in the Texas rural land market literature as well as expand the application of the DAG methodology in uncovering interactions in land markets.

3.Data and Analytical framework

Data

This work uses land price data provided by Texas A&M University, Real Estate Center. Since this study focuses on rural land price, only quarterly price over time is considered. A 52-year sample period from the third quarter of 1966 to the fourth quarter of 2017 is included. Figure2 shows the quarterly price per acre over time within the seven regions. Price per acre reports the average of the median prices per acre for markets segmented by size of property sold in each region. The price series appear to be non-stationary and exhibit similar patterns to each other.

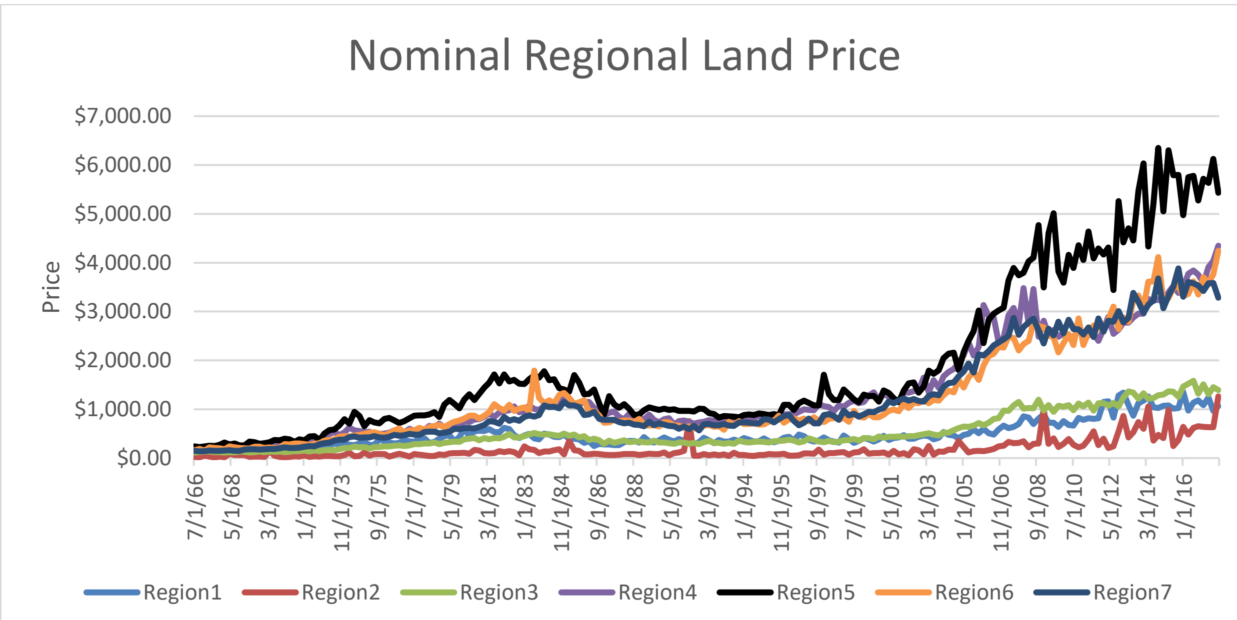

Figure2. Regional Land Price

Price in region 5, which is the Austin-Waco-Hill area ranks first among all regions, while price in Far West Texas always falls behind.Table 1 provides summary statistics of price per acre for seven regions. We can see similar trends as discussed above.

	Region1	*Region2*	*Region3*	*Region4*	*Region5*	*Region6*	*Region7*
Mean	488.38	177.55	521.68	1345.47	1870.65	1242.88	1222.16
Standard Deviation	272.57	205.56	380.90	1018.67	1611.83	995.37	1013.98
Minimum	170.00	12.92	115.00	179.38	224.63	155.50	138.00
Maximum	1354.40	1262.09	1582.22	4346.90	6349.26	4243.96	3883.05
Confidence Level(95.0%)	37.44	28.24	52.32	139.93	221.41	136.73	139.29

Table1. Summary statistics of land price per acre. Source: Generated by authors

Methodology

The directed graph technique, which represents the graphical form of causal relationships, has been widely used for non-experimental data (also known as observational data) analysis (See Chickering (2002) and Dharmasena et al. (2016) as examples.). Achievements in computer science build theoretical framework for casual relationships (Pearl, 1995; Spirtes et al., 2000). Rather than concentrating on predictability, causation also pays attention on intervention and further realization (Bessler, 2002). Such connection to intervention and manipulation gives us hope for providing appropriate forward-looking thoughts on the regional rural land markets. Below we will provide a brief discussion on the method and algorithm applied in this paper.

A directed acyclic graph(DAG) is a way to summarize causal flow among a set of variables(nodes) that is potentially causally related. The DAG contains only directed edges and does not include any cycles that start and end at the same variable. Lines with arrows($X \rightarrow Y$) indicate that variable X causes variable Y. Lines without arrows($X - Y$)indicate that X and Y are connected by information flows but the direction can not be decided by the algorithm. Two variables without connections of either lines or arrows indicate that they are not connected by any information flows. There are three types of structures in a DAG: causal chains, casual forks and inverted casual forks. Specifically, three variables X, Y and Z connected as $X \rightarrow Y \rightarrow Z$ is a causal chain. It implies that X causes Y and Y causes Z. A causal fork can be represented as $Y \leftarrow X \rightarrow Z$. It illustrates that X is a common cause of Y and Z. In the absence of X, Y and Z may become related. If X, Y and Z are related as $Y \rightarrow X \leftarrow Z$, it is called an inverted causal fork(collider). Y and Z are not related unless conditioning on X. In another words, extract information from X which is the collider will open a causality path between Y and Z.

As discussed, several machine learning algorithms exist in the literature that can generate DAGs. The Greedy Equivalence Search (GES) Algorithm developed by Chickering(2002) is used to modeling causality structure in our paper.The GES is embedded in TETRAD[5](Glymour,2016). It is a two–phase greedy search algorithm that starts a DAG with no edges, meaning that all variables are independent. In the first phase, the algorithm attempt to connect variables by adding edges. By searching stepwise, GES score all possible graphs among all equivalent classes[6] using a Bayesian

[5] TETRAD is available at: http://www.phil.cmu.edu/projects/tetrad/.
[6] Check Chickering (2002) for definition of equivalent classes.

Information Criterion (BIC)[7] and chooses the one that increases the score most for the next phase. The second phase begins by removing or reversing a single edge. The process terminates when no further deletions or reversals improves the score.

Besides assumptions reviewed in GES algorithm, a bunch of constraints can be imposed on search to generate simple but meaningful graph. Take our variables as example, we have time series data for seven regions in Texas rural land market. Besides current time relations, we are also interested in how prior data influences the current market. However, the data cannot distinguish between graphs $x_t \to x_{t-1}$ and $x_{t-1} \to x_t$. We need to impose background knowledge to decide the directions of arrows. In our analysis, edges from a later tier are forbidden in earlier knowledge tiers. This indicates that current data will not cause prior data in our analysis. Moreover, edges between two lags will be forbidden because we care about interrelationships at contemporaneous time period.

Structural relationships inherent to the DAGs can be derived afterward. A comprehensive discussion on estimating the corresponding structural models can be found in Bollen (1989).For simplicity, only a brief explanation will be provided. The causal structure of a structural model is specified using directed edges from the DAG. For example, the directed edge in $X \to Y$indicates that Y could be represented by a function of X(i.e. $Y = \beta X + \mu$). Notably, we assume that the causal model is linear (for simplicity) with Gaussian errors (since the GES algorithm requires that the underlying data generating process to be Gaussian). Coefficients and residual variances are estimated using a multiple linear regression. The number of partial effects to be estimated is equal to the number of edges in the graph.

There are two criteria which play an important role in identifying parameters. The first is front-door criteria. Suppose we are interested in the causal relationship between X and Y as showed in Figure3 below. The set of variables Q meet the front-door criteria if: (1) Q blocks all paths from X to Y, (2) there are no unblocked back-door paths from X to Q, and (3) all back-door paths from Q to Y are blocked by X.

[7] The BIC (B(G,D)) is represented as:$lnp\left(D\middle|\hat{\theta}, G^h\right) - \frac{d}{2}\ln m$, where $\hat{\theta}$ is the maximum-likelihood estimate, d is the number of free parameters of graph G, m is number of observations in data, D. So the BIC function wants to look for a tradeoff between fit given by $lnp\left(D\middle|\hat{\theta}, G^h\right)$ and parsimony given by $-\frac{d}{2}\ln m$.

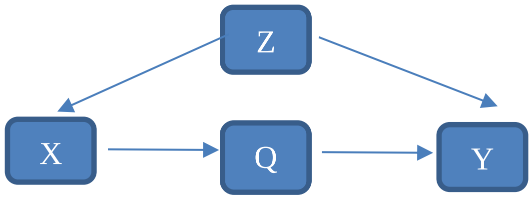

Figure3. Front-Door Criteria

To block the front-door path, one needs two steps. First, regress Y on Q and X (i.e., $Y = \beta_0 + \beta_1 X + \beta_2 Q + \mu$) to get an estimate of $\frac{\delta Y}{\delta Q}$. Second, regress Q on X (i.e., $Q = \beta_0 + \beta_1 X + \mu$) to get an estimate of $\frac{\delta Q}{\delta X}$.The unbiased and consistent estimate of $\frac{\delta Y}{\delta X}$ can be found by multiplying $\frac{\delta Y}{\delta Q}$by $\frac{\delta Q}{\delta X}$.

The second is the back-door critera. Suppose we are interested in the causal relationship between X and Y described in Figure4. The variables Z satisfy the back-door criteria if: (1) no variables in Z are descendants of X and (2) Z blocks every path between X and Y that contains an arrow into X. In order to block the information flow from X to Y via back-door path, we run a regression of Y on X and Z (i.e. $Y = \beta_0 + \beta_1 X + \beta_2 Z + \mu$). The conditioning on Z will block information flow via the back-door and provide and unbiased and consistent estimate of $\frac{\delta Y}{\delta X}$.

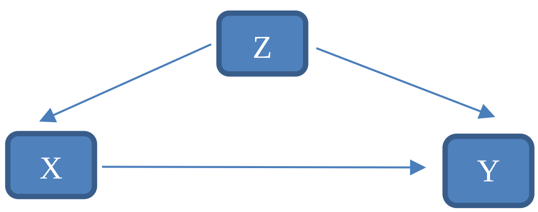

Figure4. Back-Door Critera

4. Results

Initially, a DAG constructed using time series variables is provided. Second, we present Markov Blankets for selected regions. Third, we derive causal structural models based on the DAG developed in second part.

Our variables of interest include quarterly price per acre for all seven regions and their one and two period lags. This means we have 21 variables in total. To speed up the estimation and generate edges with correct direction, we impose temporal tiers and make constraints on variables interrelationship. Specifically, we do not allow any price variables at time T influence prices in previous periods. Besides, we also exclude possible interactions within past period variables since they are supposed to display same pattern in current time. As a result, the final DAG developed for 21 variables from seven rural land markets in Texas is depicted in figure5.

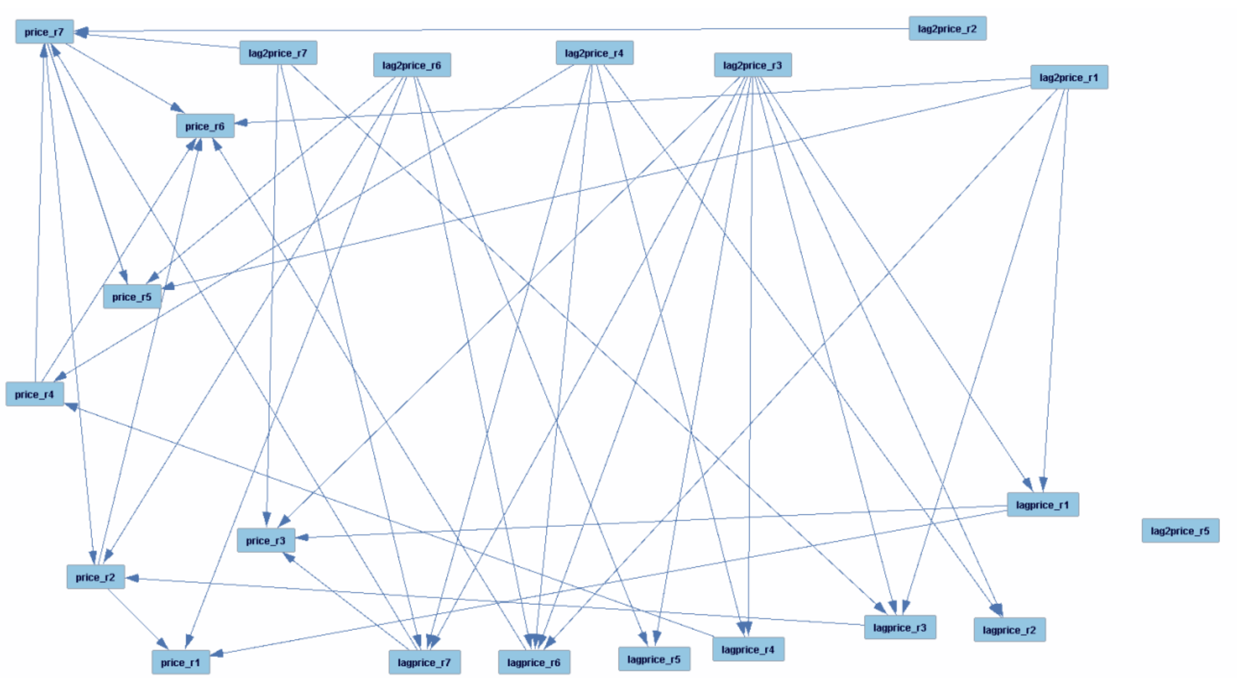

Figure5. The directed acyclic graph associated with 21 price variables using GES algorithm.

As showed, the causality structures among these variables are complicated. Furthermore, Table 2 presents estimators based on the DAG in figure. We list includes partial derivative values, standard errors associated with the partials, and associated t-statistic and p-value with respect to each directed edge under consideration.

They reveal valuable information for participants on the rural land markets. We are confidence to conclude that Texas rural land markets are not independent.

As mentioned by Senia et al.(2018), Markov blankets can help to isolate a variable from the rest of the complicated DAG. It includes only parents, children and grandparents of a single node and thus keeps the most important information for causal structure modeling. The Markov blanket of prices of seven regions with partial effects are given in figure6 to figure12. We also estimate structural models for each of the region based on the Markov blanket graph. Here we emphasis on region1,region 3 and region 5.

Figure6 shows Markov blanket of region1. Region 1 and 6 both have a lot of cropland. When harvests happen earlier in region6 for example ,pecan and cotton in the valley of region 6, it takes some time to affect the land market in region1.The number between each directed edge represents the value of the partial derivative between the variables. The *0.0000* value represent the p-value associated with the estimated partial derivative. Based on the two criteria discussed before, the structural model for region 1 is:

$$P_{1t} = \alpha_0 + \alpha_1 P_{1t-1} + \alpha_2 P_{2t} + \alpha_3 P_{6t-2} + \mu$$

where $\alpha_1 = 0.5098, \alpha_2 = 0.1770 * 0.1775 = 0.0314, \alpha_3 = 0.1007.$

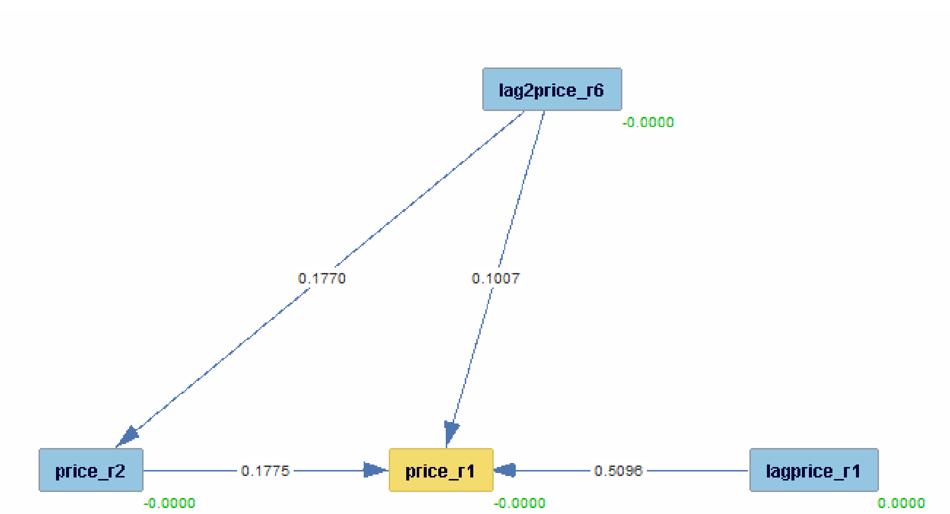

Figure6. Markov blanket of Region1 price.

Figure8 presents Markov blanket of region3. Besides its own lag, price of region 1(one lag) and region 7(one and two lags) also contribute to the price of region 3. Region7 has a very positive effect on region3. The possible explanation could be the sale of recreation lands. When buyers in region7 realize it might be cheaper to buy a larger land in nearby region, they will drive to do so. Therefore, the structural model for region3 is:

$$P_{3t} = \alpha_0 + \alpha_1 P_{7t-1} + \alpha_2 P_{1t-1} + \alpha_3 P_{3t-2} + \alpha_4 P_{7t-2} + \mu$$

where $\alpha_1 = 0.7762 * 0.1183 = 0.0918, \alpha_2 = 0.6811 * 0.2110 = 0.1437,\ \alpha_3 = 0.3093,\ \alpha_4 = 0.0925$.

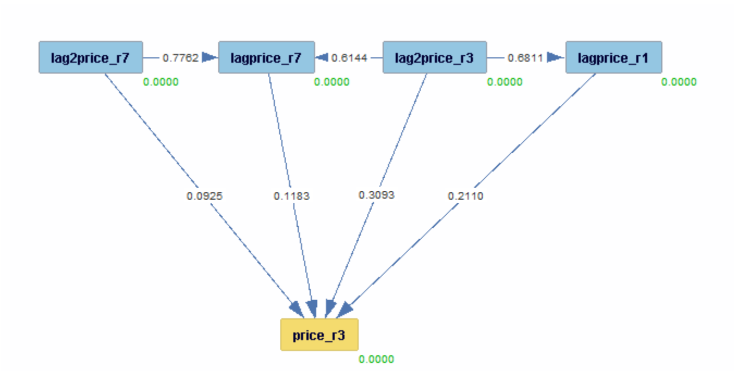

Figure8. Markov blanket of Region3 price.

Figure10 gives Markov blanket of region5 and it indicates that region1 (two lags), region6 (two lags) as well as region7 cause price change in region 5. The corresponding structural model therefore is:

$$P_{5t} = \alpha_0 + \alpha_1 P_{7t} + \alpha_2 P_{1t-2} + \alpha_3 P_{6t-2} + \mu$$

where $\alpha_1 = 0.6228$, $\alpha_2 = 0.8206$, $\alpha_3 = 0.7908$.

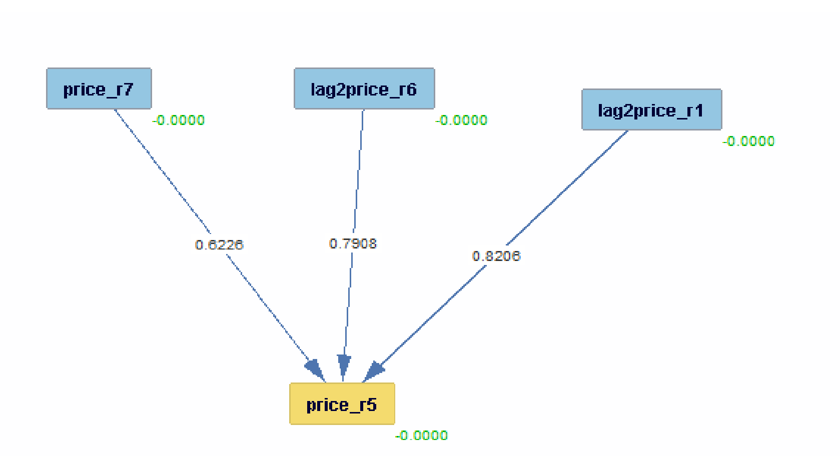

Figure10. Markov blanket of Region5 price.

Actually, all of the partial effects estimated in our DAG are positive, indicating that an increase in price in one region will inevitably boom the rural land market else where across the state.

5. Conclusions

To summarize, the goal of the research is to investigate the interrelationships between different rural land markets in Texas. Existing rural land studies place emphasis on the valuation of farmland and factors that influence it. The lack of sound method as well as specialized research investigating Texas rural land markets calls for a more comprehensive study. By applying a novel approach that inferring causal relationships from observational data in the absence of controlled experiments, this paper provides new findings for Texas rural land markets. With a strong database from Real Estate Center, our study also bridges the gap in the literature and sheds light on causal structural modeling for the land market.

Seven regional rural land markets in Texas are considered. We construct a directed acyclic graph (DAG) based on 21 variables (per acre from seven regions and lags of the price). Prior knowledge that constrains the final graph is imposed on the structure. From this DAG, we are able to estimate causality relationships and partial effects for the edges, which allowed for predicting rural land market in Texas. In particular, we centered attention on uncovering causality structures separately for each region. We justify our hypothesis that rural land markets in Texas are not independent. Therefore, besides various local factors that influence regional land price, we should take into account interrelationships among regions to get a whole picture of what is going on in the market. Additionally, the importance of understanding these causality relationships is apparent. If analysts want to model factors affecting a given market but conditioning on the wrong variable, an important path may be blocked. Our work stimulates future research to where one may want investigate the dynamic behavior of other important variables through alternative modeling approaches (Dharmasena et al.,2016).

We conclude our paper by highlighting a few limitations of the study. First, the GES algorithm applied here assumes an normal distribution of the data while the price trends showed in our data do not satisfy the assumption. As a result, moving beyond such algorithm and employing a non-Gaussian algorithm such as LiNGAM may be helpful (Shimizu et al., 2006). Second, price here is nominal price that does not eliminate inflation. A real price analysis is also a necessary direction. Third, the results reported here are dependent on the quality of the input data. Using only price in our study may take the risk of losing too much information. The study can be augmented by including more latent variables from the market.

Appendix

To	From	Partial value	Std error	t-stat	p-value
Region3	Region7(t-1)	0.1183	0.025	4.73	0.0000
Region7	Region7(t-1)	0.3765	0.0623	6.0391	0.0000
Region4	Region4(t-2)	0.5337	0.061	8.815	0.0000
Region5	Region1(t-2)	0.8206	0.1806	4.5441	0.0000
Region3(t-1)	Region3(t-2)	0.3245	0.0598	5.4233	0.0000
Region6	Region6(t-1)	0.3408	0.0529	6.4471	0.0000
Region3	Region3(t-2)	0.3093	0.0575	5.3791	0.0000
Region1	Region1(t-1)	0.5096	0.0562	9.0661	0.0000
Region5(t-1)	Region6(t-2)	0.9667	0.1208	8.002	0.0000
Region3	Region7(t-2)	0.0925	0.0264	3.5061	0.0006
Region1(t-1)	Region1(t-2)	0.5200	0.0588	8.8392	0.0000
Region6	Region1(t-2)	0.6220	0.0965	6.4423	0.0000
Region7	Region4	0.1952	0.0397	4.9183	0.0000
Region2	Region6(t-2)	0.2436	0.0455	5.348	0.0000
Region6(t-1)	Region6(t-2)	0.4106	0.0639	6.4245	0.0000
Region4	Region4(t-1)	0.4482	0.0601	8.0215	0.0000
Region7(t-1)	Region7(t-2)	0.4420	0.0679	6.5139	0.0000
Region6	Region7	0.2939	0.0644	4.5628	0.0000
Region6	Region2	0.2970	0.0745	3.9886	0.0001
Region6(t-1)	Region3(t-2)	0.6688	0.1627	4.1119	0.0001
Region7(t-1)	Region4(t-2)	0.3072	0.0455	6.7443	0.0000
Region6(t-1)	Region4(t-2)	0.2083	0.0414	5.0354	0.0000
Region3(t-1)	Region1(t-2)	0.2487	0.035	7.1027	0.0000
Region1	Region2	0.1775	0.047	3.7791	0.0002
Region7(t-1)	Region3(t-2)	0.7096	0.1248	5.6854	0.0000
Region1(t-1)	Region3(t-2)	0.3301	0.0425	7.7711	0.0000
Region5	Region6(t-2)	0.7908	0.0985	8.0262	0.0000
Region7	Region2(t-2)	0.2732	0.066	4.1427	0.0001
Region7	Region7(t-2)	0.3956	0.0612	6.4693	0.0000
Region2(t-2)	Region3(t-2)	0.6762	0.0754	8.9621	0.0000
Region5(t-1)	Region3(t-2)	1.7349	0.3123	5.5547	0.0000
Region2(t-1)	Region4(t-2)	-0.0982	0.0285	-3.4435	0.0007
Region1	Region6(t-2)	0.1007	0.0176	5.7288	0.0000
Region6	Region4	0.1533	0.047	3.2639	0.0013
Region4(t-1)	Region4(t-2)	0.7591	0.0436	17.415	0.0000
Region2	Region7	-0.2368	0.0433	-6.098	0.0000
Region3(t-1)	Region7(t-2)	0.1926	0.0193	9.9712	0.0000
Region5	Region7	0.6226	0.0861	7.2316	0.0000
Region6(t-1)	Region1(t-2)	0.4910	0.1044	4.7036	0.0000
Region3	Region1(t-1)	0.2110	0.0336	6.2826	0.0000
Region4(t-1)	Region3(t-2)	0.6606	0.1154	5.7267	0.0000
Region2	Region3(t-1)	0.5339	0.1282	4.1637	0.0000

Table2. Parameter estimates (partial value) for each edge and their associated significance.

Source: calculated by authors

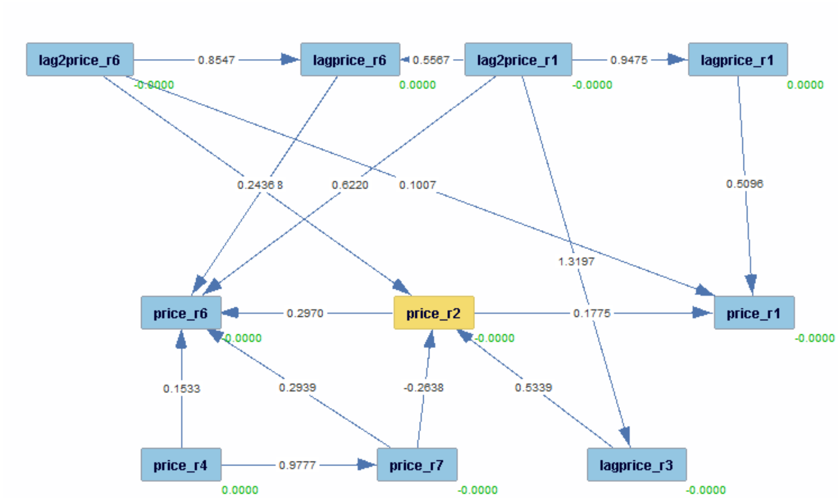

Figure7. Markov blanket of Region2 price.

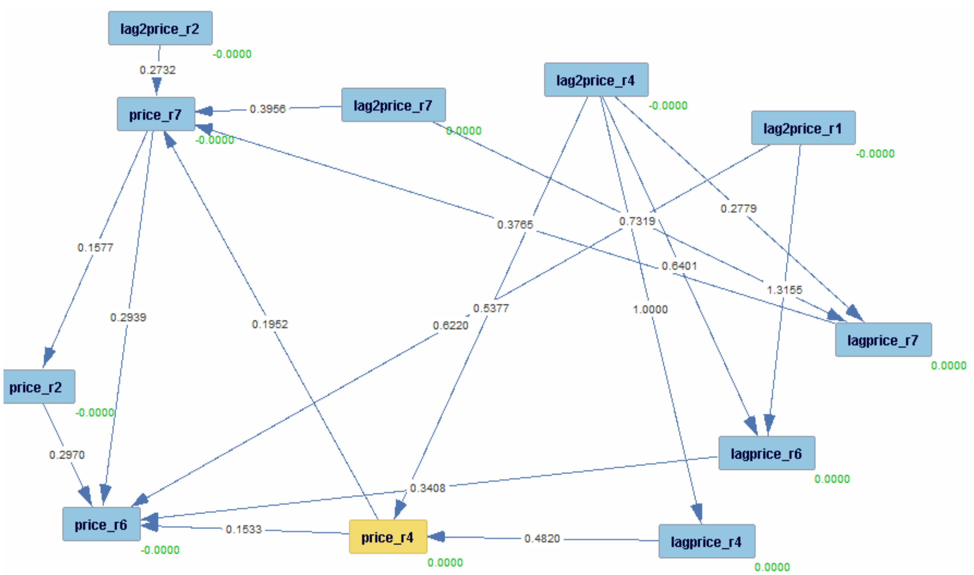

Figure9. Markov blanket of Region4 price.

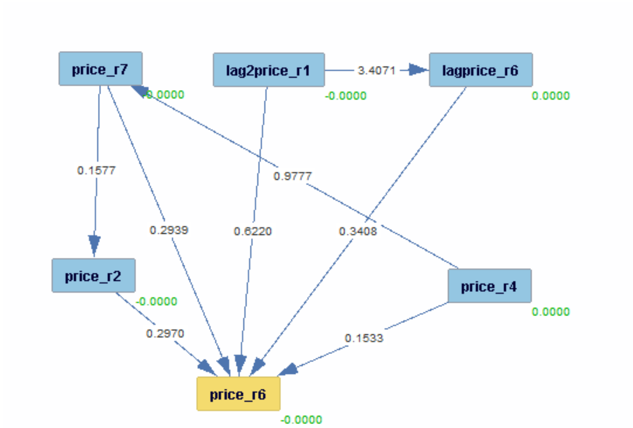

Figure11. Markov blanket of Region6 price.

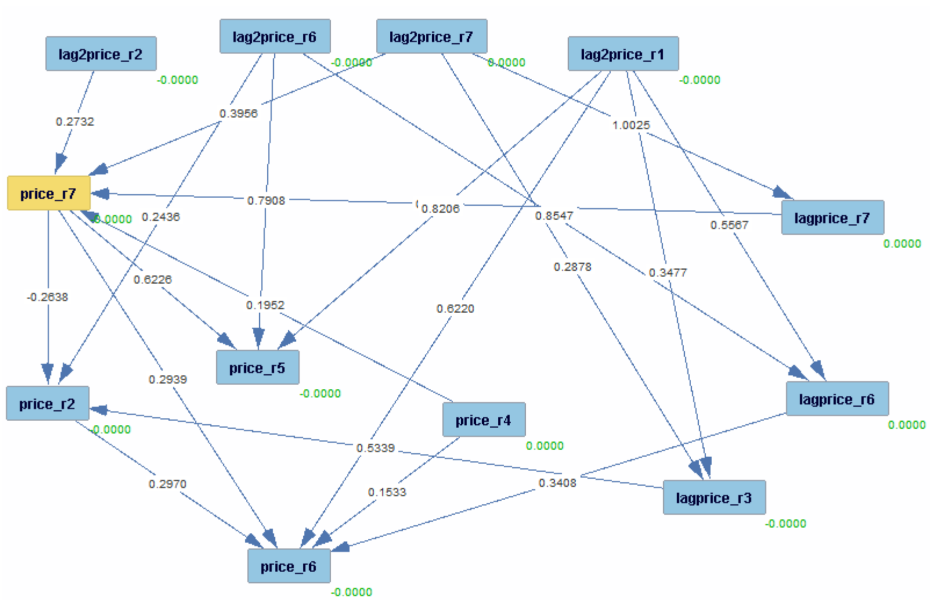

Figure12. Markov blanket of Region7 price

Reference

Alexander, C. and Barrow, M., 1994. Seasonality and cointegration of regional house prices in the UK. *Urban Studies*,*31*(10), pp.1667-1689.

Bessler, D.A., 2003. On world poverty: Its causes and effects. *Food and Agricultural Organization (FAO) of the United Nations, Research Bulletin, Rome*.

Bessler, D.A. and Akleman, D.G., 1998. Farm prices, retail prices, and directed graphs: results for pork and beef. *American Journal of Agricultural Economics*,*80*(5), pp.1144-1149.

Bollen, K.A., 2014.*Structural equations with latent variables*(Vol. 210). John Wiley & Sons.

Chickering, D.M., 2002. Optimal structure identification with greedy search. *Journal of machine learning research*,*3*(Nov), pp.507-554.

Dharmasena, S., Bessler, D.A. and Capps Jr, O., 2016. Food environment in the United States as a complex economic system. *Food Policy*,*61*, pp.163-175.

Fang, L. and Bessler, D.A., 2017. Stock returns and interest rates in China: the prequential approach. *Applied Economics*,*49*(53), pp.5412-5425.

Faubion, J.F., 1971.*Analysis of the effects of consumptive uses on the value of rural land* (Doctoral dissertation, Texas A&M University).

Gilliland, C.E., Gunadekar, A., Wiehe, K. and Whitmore, S., 2010.*Characteristics of Texas Land Markets: A Regional Analysis*. Real Estate Center.

Gilliland, C.E., 2012. Buying Rural Land in Texas. Texas A&M University Press.

Glymour, C., R. Scheines, P. Spirtes, and J. Ramsey. 2016. The TETRAD Project: Causal Models and Statistical Data. Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA, http://www.phil.cmu.edu/tetrad/current.html.

Holmes, M.J. and Grimes, A., 2008. Is there long-run convergence among regional house prices in the UK?.*Urban Studies*,*45*(8), pp.1531-1544.

Jones, D.W., Ring, J.K., Jones, J.C., Watson, K., Parvin, D.W. and Munn, I.A., 2006. Land valuation increases from recreational opportunity: a study of Mississippi rural land sales. *Proceedings of the Annual Southeastern Association of Fish and Wildlife Agencies*, *60*, pp.49-53.

Kennedy, G.A., Henning, S.A., Vandeveer, L.R. and Dai, M., 1997. *An empirical analysis of the Louisiana rural land market* (No. 1083-2016-87308).

Lee, C.C. and Chien, M.S., 2011. Empirical modelling of regional house prices and the ripple effect. *Urban Studies*,*48*(10), pp.2029-2047.

Lee, M.T., Lee, M.L. and Lin, S.H., 2014. Trend properties, cointegration, and diffusion of presale house prices in Taiwan: Can Taipei's house prices ripple out? *Habitat International*,*44*, pp.432-441.

Luo, Z.Q., Liu, C. and Picken, D., 2007. Housing price diffusion pattern of Australia's state capital cities. *International Journal of Strategic Property Management*,*11*(4), pp.227-242.

MacDonald, R. and Taylor, M.P., 1993. Regional house prices in Britain: long-run relationships and short-run dynamics. *Scottish Journal of Political Economy*,*40*(1), pp.43-55.

Meen, G., 1999. Regional house prices and the ripple effect: a new interpretation. *Housing studies*,*14*(6), pp.733-753.

Munro, M. and Tu, Y., 1996. The dynamics of UK national and regional house prices. *Review of Urban & Regional Development Studies*,*8*(2), pp.186-201.

Pearl, J., 1995. Causal diagrams for empirical research.*Biometrika*,*82*(4), pp.669-688.

Pollakowski, H.O. and Ray, T.S., 1997. Housing price diffusion patterns at different aggregation levels: an examination of housing market efficiency. *Journal of Housing Research*,*8*(1), p.107.

Pope III, C.A., 1985. Agricultural productive and consumptive use components of rural land values in Texas. *American Journal of Agricultural Economics*,*67*(1), pp.81-86.

Rettenmaier, A.J. and Wang, Z., 2012. Regional variations in medical spending and utilization: a longitudinal analysis of US Medicare population. *Health Economics*,*21*(2), pp.67-82.

Senia, M.C., Dharmasena, S. and Todd, J.E., 2018.*A Complex Model of Consumer Food Acquisitions: Applying Machine Learning and Directed Acyclic Graphs to the National Household Food Acquisition and Purchase Survey (FoodAPS)*(No. 2015-2018-104).

Shi, S., Young, M. and Hargreaves, B., 2009. The ripple effect of local house price movements in New Zealand. *Journal of Property Research*,*26*(1), pp.1-24.

Spirtes, P., Glymour, C.N., Scheines, R. and Heckerman, D., 2000. Causation, *prediction, and search*. MIT press.

Stevenson, S., 2004. House price diffusion and inter-regional and cross-border house price dynamics. *Journal of Property Research*,*21*(4), pp.301-320.

Tsai, I.C., 2014. Ripple effect in house prices and trading volume in the UK housing market: New viewpoint and evidence. *Economic Modelling*,*40*, pp.68-75.

Yu, Y., Chen, J., Gao, T. and Yu, M., 2019. DAG-GNN: DAG Structure Learning with Graph Neural Networks. *arXiv preprint arXiv:1904.10098*.

Zhang, L., Hui, E.C. and Wen, H., 2017. The regional house prices in China: Ripple effect or differentiation. *Habitat International*,*67*, pp.118-128.

Zhang, T., Gerlowski, D. and Ford, D., 2014. Housing price variability: national and local impacts. *Applied Economics*,*46*(28), pp.3494-3502.

2016. USDA ERS. February. Accessed February 2016.

2017 USDA ERS. August. Accessed August 2017.