THE VALUE OF GOOD NEIGHBORS: A SPATIAL ANALYSIS OF THE CALIFORNIA AND WASHINGTON WINE INDUSTRIES

Nan Yang, Jill J. McCluskey, and Michael Brady
School of Economic Sciences, Washington State University, Pullman, WA 99164
Email: nan.yang@email.wsu.edu

THE VALUE OF GOOD NEIGHBORS: A SPATIAL ANALYSIS OF THE CALIFORNIA AND WASHINGTON WINE INDUSTRIES

Nan Yang, Jill J. McCluskey, and Michael Brady*

Abstract: The fact that wineries tend to cluster in certain sub-regions can be partially explained by the terroir of those areas. However, a gap in our understanding of the spatial relationships among wineries remains. In this article, winery-level data with geographic information system (GIS) coordinates are utilized to examine the spatial relationships among neighboring wineries. Spatial effects for the California and Washington wine industries are assessed by performing clustering tests based on wine prices and tasting scores. A spatial lag model is then estimated to test the hypothesis that there are positive effects from neighbors when analyzing the hedonic price equations. The regression results indicate that there exists strong and positive neighbor effect.

Key words: GIS, clustering, spatial lag model, wine

*School of Economic Sciences, Washington State University, Pullman, WA 99164

Introduction

One bad wine in the valley is bad for every winery in the valley. One good wine in the valley is good for everyone. --Robert Mondavi on the Napa Valley in the 1960s (Stiler, 2007).

When one examines a map with points indicating winery locations in California and Washington State, there is an interesting phenomenon. Wineries are intensively located in some of the areas but almost none of them in others. In other words, most of the wineries choose to locate close to each other. The obvious reason for this location pattern is probably because of geographic features as defined by American Viticultural Area: the *terroir* of some regions is more suitable for grape growing. Therefore, wineries prefer to select a location that can explore this resource advantage. However, the reason that wineries do not evenly distribute within grape growing region but choose to cluster together cannot be well explained by *terroir*. Therefore, a research question comes out naturally. Do wineries benefit from choosing locations that are in close proximity to high reputation neighbor wineries?

Many studies on wine markets have considered location as factor, but geographic clustering and neighborhood effects (micro-level interactions among wineries) in wine industry have not been fully analyzed. The idea behind Tobler's "First Law of Geography," is that everything is related to everything else, but close things are more related than distance things (Tobler 1970). Following this idea, the influence from neighbors may be quite important to a winery's product prices.

Possible reasons for winery concentrations can come from both the producer and consumer points of view. First, from production side, spatial heteroscedasticity and spatial dependence can be the reasons. Spatial heteroscedasticity refers to the *terroir* of a sub-region. It is an exogenous factor. Since a winery and its neighbors share the same grape growing conditions, it is possible that their products exhibit some degree of similarity in terms of their characteristics. This may result in similar tasting scores from wine experts, and many previous studies indicate that score is the most important factor in determining a wine's price. Spatial dependence represents the spillover effects of reputation and management among wineries, which are located in close proximity. This is an endogenous factor: nearby wineries usually are located in the same appellation, which serves as a reputation signal to the market. Also, knowledge about how to make wine is easier to be communicated. Therefore, closeby wineries are able to and willing to charge a similar price for their products.

Second, from consumer's side, perceptions about wines coming from the same area tend to be the similar, so it is more likely for consumer to be willing to pay similar price for wines from the same micro-region. For example, if a winery with a reputation for producing extremely high quality wine is located close to another winery, consumers might consider the neighboring winery to also have high quality products. Therefore, we cannot ignore the impacts from neighbors of a winery.

According to Can (1998), spatial analysis is usually divided into two stages. The first stage is the exploratory spatial data analysis (ESDA) stage or spatial pattern identification stage, which concerns description rather than explanation. The second stage is called confirmatory data

4

analysis (CDA) stage, which involves modeling the impact of spatial structure on behavior and outcomes in addition to economic considerations. In this study, we will follow this framework to analyze the spatial effects for Washington and California wine industries by first performing clustering tests based on prices and tasting scores and then formally measuring neighborhood effects via a spatial lag model.

Porter (2000) is the seminal article in the literature on geographic clustering analyses. Porter explains the literature and methods for clusters, or geographic concentrations of interconnected companies, and their role in competition and other implications. He argues that clusters represent a new way of thinking about economics, and they necessitate new roles for companies, government, and other institutions in enhancing competitiveness. For the identification of clusters, he indicates that the ultimate determining factors of a cluster are the strength of the "spillovers." The geographic scope of a cluster is related to distance, and informational, transactional and other efficiencies occur over the cluster. He also argues that all existing and emerging clusters deserve attention. From an empirical point of view, Jaffe *et al*. (1993) uses patent citations to test whether the knowledge spillovers are geographically localized. They find that localization does exist and it slowly fades over time. However, no previous studies have focused on geographical clustering of wine and its effects. Other types of clustering analysis has been a major topic for some studies. Costanigro *et al*. (2009) identify wine segments for Washington and California wines with a procedure called local polynomial regression clustering, which is clustering by local regression coefficients, but Geographic Information System (GIS) data was not used. Kaye-Blake *et al*. (2007) utilize cluster analysis on potential market segments

for genetically modified food. However, their clusters are based on survey responses.

For modeling spatial effects, Anselin (1999) summarizes the foundation and regression issues of spatial models. In the applications of spatial model, Wu and Hendrick (2009) estimate a spatial lag reaction function for property tax rate of Florida municipal governments in 2000 and 2004 and compare model fitness as well as results from different specifications of spatial relationships. Garretsen and Peeters (2009) test the relevance of spatial linkages for Dutch (outbound) foreign direct investment (FDI). They estimate a spatial lag model for Dutch FDI to 18 host countries and find that third-country effects matter.

GIS can be a helpful and powerful tool in spatial relationship studies. As indicated by Can (1998), GIS enables the researcher to organize, visualize, and analyze data in a map form, provides the medium for the integration of multiple geographical data sets, and gives analytical support for spatial data analysis by providing explicit information of spatial relationships. Wallsten (2001) applies GIS and firm-level data to explore agglomeration and spillovers at the firm level over discrete distances. He finds that the number of other firms participating in Small Business Innovation Research (SBIR) program within a fraction of a mile predicted whether a firm wins awards.

In this article, winery-level GIS data is collected to describe the spatial relationships among wineries for California and Washington State wine industries. First, we conduct formal statistical tests to decide whether there exists geographic winery clusters based on price and tasting score. Second, a spatial lag model is applied to test the hypothesis that there is positive effect from neighbors when analyzing the hedonic relationships among price and other factors. Analyses are

done for both California and Washington State.

Spatial analyses of California and Washington State wine industry can improve the understanding of the economic relationships among wine's price and product's attributes. All previous studies about Washington and California wine industries either ignore the spatial autocorrelations among wineries or wine regions, or treat them as nuisance and incorporate them into the error structure of the regression model. However, there is a high possibility that spatial autocorrelations (spatial effects), which may be the results of spatial interaction processes, externalities, spillover and so on, is significantly present among Washington and California wineries. If we ignore the spatial nature of the data, it may lead to biased or inefficient estimates and misleading inference (Anselin, 1988). Consequently, this research can help to look for a more appropriate econometric model to describe the relationships among price and other characteristics of Washington and California red wines, when considering the spatial effects from the distribution of hundreds of existing wineries.

The rest of this article proceeds as following: First, we will describe the data used in this study. Second, we will introduce the econometric methods and model applied in this study, and the statistics test for geographic clustering. Following these are the results and discussion. Last, we offer conclusions.

Data

The data set consists of winery-level data from two States: Washington and California.

For each observation, information about price, rating score, case, year of aging, vintage and production region is collected from *Wine Spectator* magazine (online version). Since the observed unit in this study is individual winery, the above variables of price, score, case and age are averaged across grape varieties[1] and vintages[2] for every winery in our data set. Indicator variables are used to denote the winery's production area, representing collective reputations. The regions for California wines include Napa Valley, Bay Area, Sonoma, South Coast, Carneros, Sierra-Foothills, Mendocino and other California. In Washington, they are Columbia Valley, Yakima Valley, Walla Walla Valley, Puget Sound and other Washington.

For Washington State, information about 79 wineries is gathered. For California, there are 876 wineries in our data set. Table 1 reports the descriptive summary of non-binary variables in our data set and Table 2 provides short descriptions and abbreviations of all variables used in the empirical analyses.

To describe the spatial property of each winery, we incorporate GIS data into our study. Each winery contains a name, street address, city, state and zip code. The street address in the data allows us to recover each winery's exact longitude and latitude coordinates by geocoding address in GIS program[3]. Following geocoding, we can obtain an accurate understanding of almost any

---

[1] This study only takes red wines into concern; the grape varieties are Zinfandel, Pinot Noir, Cabernet Sauvignon, Merlot, Syrah grapes, and wines made from blending of different varieties (non-varietals).

[2] The vintages are from 1991 to 2000.

[3] The data set originally contains 137 wineries for Washington and 1195 for California. However, due to the difficulty of finding street address (e.g. some wineries only provide P.O. box or can only locate to city), and GIS's limited ability

spatial relationship among wineries in our data set, such as pairwise distances between any two wineries and the nearest *K* neighbors for any selected winery. Also, we are able to obtain a visual understanding about the spatial distribution of wineries in both Washington and California. Figures 1 and 2 are winery distribution for Washington and California, respectively.

About the spatial information of our data set, two things need to be mentioned. First, wine Spectator is the only source of our data set. We only include wineries whose wines are listed in Wine Spectator. Second, among all the wineries, in Washington State, 10.13% of them are estate wineries and 4.57% are estate wineries in California. Only these wineries use their own grapes to produce wine instead of buying them from external growers. The coordinate of each winery is where the producing processes take place.

Method and Model

There are two distinct ways to model spatial dependence: as an additional regressor in the form of a spatially lagged dependent variable (Wy), or in the error structure (E [$\varepsilon_i \varepsilon_j$]≠0). The former one is referred as a spatial lag model and has the form of $y = \rho Wy + X\beta + \varepsilon$, and the later one is usually called spatial error model with the expression $y = X\beta + \varepsilon$ and $\varepsilon = \lambda W\varepsilon + u$. The choice of the model depends on the research interest. When the focus of interest is to assess the existence and strength of spatial interaction, the spatial lag model is more appropriate, since it

to locate some addresses, only 79 wineries from Washington and 876 wineries from California can be applied in the study.

interprets the spatial dependence in a substantive form. However, when the concern is to correct

the potentially biasing influence of the spatial autocorrelation due to spatial data, the spatial error

model is appropriate to meet the goal (Anselin 1988).

The prior objective in this study is to model spatial effects in California and

Washington State wine industries. Therefore, it is necessary to include a specific term of

neighborhood effect in the explanatory variables, and the spatial lag model will be a reasonable

choice.

Regarding other explanatory variables, we choose to include factors showing

significant effects in many previous hedonic analyses of wine. Therefore, the formal expression of

our spatial lag model is

$$
\begin{aligned}
\mathrm{f(Price)} = {} & \rho W(f(\mathrm{Price})) + \beta_0 + \beta_1(\mathrm{Score}) + \beta_2(\mathrm{Case}) + \beta_3(\mathrm{Age}) \\
& + \sum_{i=1}^{j} \beta_{3+i}(\mathrm{Region}_i) + \varepsilon
\end{aligned}
\tag{1}
$$

Where ρ is a spatial autoregressive coefficient, W is the spatial weight matrix that will be specified

later. *Score* is the rating score from *Wine Spectator* magazine, *Case* is the number of cases

produced by the winery, and *Age* represents years of aging before commercialization. All of these

three variables are averaged values for the particular winery across the observation period. *Region*

tells us the place of production and each area is represented by an indicator variable. For

Washington State, there are four regions (*j* is equal to four). The regions are the Columbia Valley,

Yakima Valley, Walla Walla Valley and Puget Sound. For California, there are seven macro

regions ($j = 7$). They are Napa Valley, Bay Area, Sonoma, South Coast, Carneros, Sierra-Foothills

and Mendocino are the seven production regions. Since for both states, the region variables

exclude Other Washington or Other California, parameter before each of them indicates the

difference between wine from this area and a generic Washington or California wine.

The form of the dependent variable *f(Price)* is determined by a Box-Cox

transformation. For Washington State, we use *ln(Price)* as our final dependent variable in the

regression, while for California, the best transformation is *Price$^{-0.25}$*. Equation (1) is estimated via

spatial econometrics method.[4]

The specification of spatial weight matrix in spatial econometric analysis is important

and influential to regression results. In previous studies, Frizado *et al*. (2009) emphasize the

sensitivity of spatial weights matrix selection to the cluster identification results by Local Moran's

and Getis-Ord Gi when concerning U.S. county size. They conclude that the selection of spatial

weighting methodology should depend on the study's purpose, the distribution of county sizes, and

the industry being studied. Also, Anselin (1999) points out that the elements of the weights matrix

are non-stochastic and exogenous to the model. Typically, they are based on the geographic

arrangement of the observations, or contiguity. Several forms of spatial weights are analyzed in

literature, such as inverse distance or inverse distance squared (Anselin, 1980), structure of a social

network (Doreian, 1980), economic distance (Case, Rosen and Hines,1993) and K nearest

neighbors (Pinkse and Slade, 1998).

However, the specification of spatial weights is not arbitrary. The range of

---

[4] See Anselin (1988)

dependence allowed by the structure of W must be constrained. Therefore, the key question in

every spatial econometric analysis is how to define the range of the neighborhood. Intuitively, if

the units all belong to one cluster, then distance decay will be a reasonable choice of spatial

weights because it treats all units as neighbors. However, when units are distributed as several

"hot spots" in space, only consider distance weight will not be a good candidate. Since treating a

far-away point, which belongs to another cluster, as neighbor does not make sense. Also, to avoid

confusing the exogeneity of weights, deriving weights geographically is more appropriate.

Therefore, based on the geographic distribution of California and Washington State

wineries, we select K nearest neighbors[5] as the structure of our spatial weight matrix. As the

empirical standard of model selection, we also compare Akaike's Information Criterion (AIC) of

models with different spatial weight matrix and find that the *K* nearest neighbors structure that

results in the best AIC value. The AIC measurements have also been applied in a number of

spatial analyses as mentioned in Anselin (1988, page 247).

Clustering Test (Global Moran's I)

Before proceeding to the spatial econometric analysis, it is necessary to get an

approximate idea of how well the geographic connection is among wineries. Formal measurement

of trends in spatial pattern can be accomplished by spatial association (or autocorrelation) statistics.

The most common one to identify geographic cluster is Moran's I statistic, which is derived from a

---

[5] Only the nearest K wineries are considered to have influence on the interest winery.

statistic developed by Moran (1948, 1950a, 1950b). Also there are Geary's c, Gamma, Gi and Gi[*] as summarized by Anselin (1998). In the literature, geographic cluster analysis is widely applied in housing market. Anselin and Can (1995) use an exploratory spatial approach to the examination of spatial structure in 1990 mortgage originations for Dade County, Florida. They apply local spatial association statistics to identify areas that exhibit statistically significant clustering of high and low levels of mortgage activity (i.e., "hot spots").

In this study, we are interested in testing the general connection among all wineries from a State. Therefore, to evaluate whether wineries' spatial distribution pattern expresses clustered, dispersed, or random, Global Moran's I statistics is appropriate. Also, geographic connection can be based on many aspects; the ones we choose in this study are wine's price and rating score.

Global Moran' I statistics is defined as

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

(2)

where N is the number of spatial units indexed by i and j; X is the variable of interest, here is wine's price premium; $\bar{X}$ is the mean of X, or the mean of price premium; and w_{ij} is a matrix of spatial weights, which are defined by K nearest neighbors criteria.

Values of Global Moran's I range from -1 (indicating perfect dispersion) to +1 (perfect clustering). Inference for Global Moran's I is based on a normal approximation. The Z-score

value is calculated to decide whether to reject the null hypothesis that there is no spatial clustering. If the threshold significance level is set at 0.05, then a Z score would have to be less than −1.96 or greater than 1.96 to be statistically significant.

For this study, the results from Global Moran's I tests of price and score for Washington State and California are showed in Table 3 and 4. The *K* nearest neighbors spatial weight matrix is in use, and *K* is from 1 to 5 for Washington and between the range of 1 to 65 for California. The way to decide number of *K* will be discussed later. From the results, no matter how many neighbored wineries are considered, both price and score exhibit positive clustering distribution at the global level. Also, by comparing the Moran's I values, price clustering is stronger than score clustering.

A Global Moran's I test can be considered as a spatial pattern identification test. From the results, we can obtain a general understanding about the degree of spatial connection among winery's product price and quality (represent by tasting score) for both States.

Results and Discussions

Estimation results are presented in Table 5 and Table 6 for Washington State and California, respectively. Results are reported based on the choice of spatial weight matrix. The parameter ρ represents the degree of spatial effect among wineries. The probability values for each estimate are in parenthesis. For ease of comparison, we also provide estimation results from hedonic model without spatial lag term. They are presented in the last column of each table.

14

In the last row of Table 5 and 6, we compare the models based on their AIC. The model with the smallest AIC is highlighted. Also, three statistical tests (Wald, Likelihood Ratio and Lagrange Multiplier tests) are conducted to evaluate the hypothesis that there is no significant spatial effect among wineries within a state.

Washington State

For Washington State, we consider the nearest neighbors to be the closest 1 to 5 wineries. Therefore, K is less or equal to 5 in the spatial weight matrix. The reason for considering K in this range is that the AIC values a minimum point at $K = 3$ in this range. Three wineries represent about 4% of the total wineries that are listed for Washington state in the *Wine Spectator* ratings data base. There are only 79 observations for Washington State. From the results, we can see that regardless how many wineries are considered to be neighbors or potential candidates for spatial effects (from 1 to 5), according to the Wald, Likelihood Ratio, and Lagrange Multiplier tests, ρ is highly significantly different from zero and has a positive sign[6]. Since from the model specification, ρ is the parameter describing spatial correlation, this result indicates that neighbors do have significant and positive effect on a winery's own product price. Therefore, good neighbors can have beneficial impact to a winery, or we may say that there are positive neighborhood effects among Washington State wineries. This finding is consistent with positive

---

[6] Except "nearest 1 neighbor" is significant at 0.1 level or insignificant in LM test, others are at 0.01 level.

spillover theory, and it can be important to potential investors who are interested in developing new wineries in Washington State.

The AIC is used as the criteria for model selection. For Washington State, the three nearest neighbors spatial structure performs best with the smallest AIC of 15.6105. Consequently, in the following discussion, the results we refer to are from this model. For hedonic regression estimates, we obtain similar results as previous studies. *Score* has significant positive effect on price, indicating that expert evaluations have important influence on wine price. *Case* has significant negative impact on price, which is consistent with supply-demand theory: massive production may reduce price. *Age* affects price positively, which means that as the year of aging increasing, wine's value increases. For region dummy variables, all of them except *Columbia Valley*[7] are insignificant. Therefore, for the most cases, regional difference is not obviously present in Washington State wine industry, and this is probably the reason why people usually do not refer to micro wine production region for Washington State as do when they refer to California wine appellations.

Comparing the estimation results from the spatial model to hedonic regression results without considering spatial effects (Column 4 and Column 7 in Table 5); we see that there are not many differences. However, it is still necessary to consider spatial effects when conducting hedonic analyses, because from the comparison of AIC[8], model with spatial term is an improvement of simply hedonic regression.

---

[7] Wines from Columbia Valley generate discount comparing to other Washington red wines.

[8] The AIC of hedonic model without spatial lag for Washington is 19.73.

<u>California</u>

For California, in the spatial weight matrices, we consider the number of significant nearest neighbors K to be between 1 and 65. We expect for the number of wineries considered as neighbors is much greater than that of Washington State because California has more wineries and the distance between wineries are generally smaller. In our data set, there are 876 wineries in the *Wine Spectator* ratings data base for California. Therefore, the number of wineries within a given area is greater compared to Washington, and a winery is likely to have more neighbors that may have potential spatial dependence. We find that the AIC statistic decreases and reaches its minimum when 35 wineries are considered as neighbors, which is also about 4% of the total wineries. After that, the value of AIC is fairly stable.

From the estimates of spatial lag parameter ρ together with the Wald, LR and LM tests, we find that good neighbors have significant and positive effects on winery's product price in California. Further, if we compare this positive spatial effect to its counterpart in Washington State, it shows that wineries in California may experience a more apparent neighborhood impact, because the probability values of spatial term estimates are all close to zero. This is the case because California has a much longer history and more established reputation for producing wine. The nature resources of grape growing are almost fully explored by wine investors, the intensity of wineries within a small sub-region is much greater than that of Washington State. Therefore, the

connections among those close wineries are stronger due to this smaller distance between each others.

Among the K nearest neighbors spatial models we estimated for California, the model considering 35 nearest neighbors has the best AIC of -3173.595. Comparing this "best" nearest neighbor number to Washington State, where smallest AIC comes from the 3 nearest neighbors model, we may conclude that the good neighbor impact is more inclusive for California since a lot of surrounding neighbors of a winery can provide potential benefits. The following results discussions are based this model.

Since the dependent variable is the -0.25 power transformation of price, a negative sign for parameter estimate indicates a positive marginal effect on price. From the results, we can see that there are similar parameter estimates for the common factors on wine price in California as for Washington State: *Score* and *Age* have positive effect, while *Case* affects price negatively. All of the three variables are significant. However, for macro regions, except *SierraFoothills*, all other regions have significant price premium comparing to generic California red wines. This finding shows that micro region differences are present in California and is consistent with consumer's perception of the area.

Also, by comparing the spatial model to usual hedonic model, parameter estimates are not so much different. However, the spatial model is better according to AIC criteria. The AIC of the hedonic model without spatial term for California is -3124.117, which is greater than the AIC for all other spatial models in this study.

Two interesting points deserve more deep discussions: (1) Tradeoff between price and cost, (2) Long run effects from spatial correlation. First, spatial estimations for both Washington State and California suggest that clustering exists and positively affects price. These findings support the spillover effects from knowledge and reputation. However, for an entrepreneur who wants to start a new winery, this does not necessarily mean choosing the location right next to a high reputation winery is the best strategy and will generate maximum profit. There is a tradeoff between higher product price and greater cost. The land next to high reputation winery may have the opportunity of higher wine price, but it is likely that the added value is captured by the land. On the other hand, selecting land away from neighbors may cost much less and leave the new firm money to invest in other quality-affecting production factors. This study only focuses on the market price but not producing cost. Consequently, one cannot conclude that locating nearby a high reputation winery will guarantee a greater profit.

Second, from a dynamic point of view, results from this spatial analysis are related to the evolution of reputation and quality. Since locating nearby a high reputation and high price wineries may have price advantage, besides high quality wine producers, low quality wine makers will also be attracted to this area. They produce low class wine but enjoy a higher reputation and price. Moral hazard and adverse selection problems may occur, and thus, the location may no longer be an effective signal for consumers to distinguish good wines from bad ones. In the long run, the collective reputation of the sub-region will be negatively affected. Therefore, possible

dynamic equilibrium of wine quality for the sub-region tends to be lower than the initial quality. This can be considered as a by-product of the positive spatial effects among close wineries.

Conclusion

This article analyzes spatial effects of winery locations on wine price for both California and Washington State. We first located each winery in our data set accurately on the U.S. map to obtain a visual understanding of winery distribution. Since the precise longitude and latitude are available with GIS software, we can identify "neighbors" for each winery, either by a distance or nearest criterion. We use the "K nearest neighbors" approach as the standard to describe the neighborhood of wineries for both states, since it is more appropriate to the winery distribution.

From Global Moran's I clustering test, wine price and score show significant clustering patterns. This can be the starting point of spatial analyses and confirm our hypothesis about spatial effect among wine producers. Following the statistical tests, formal models are developed for both states. Spatial econometrics methods are applied and the regression results indicate that there exists strong and positive neighborhood effect: if neighbors of a winery had price premium, it is likely that the winery also has price advantage. Therefore, we can conclude that good neighbors have important values. However, the positive spatial effect cannot guarantee maximum profit for a new wine producer who is going to locate in a high price neighborhood because of the tradeoff between higher wine price and higher land cost. Also, this good neighbor

value may cause lower dynamic quality equilibrium because it will induce moral hazard and adverse selection problems.

This article is the first one to consider spatial effects of wineries in the United States. It provides a new way to apply hedonic analysis on wine price and discovers that location interactions are very important to winery's product price.

References

Anselin, L. (1988a). *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.

Anselin, L. (1998a). GIS research infrastructure for spatial analysis of real estate markets. *Journal of Housing Research* 9, 113–33.

Anselin, L. and Can, A. (1995), Spatial Effects in Models of Mortgage Origination. *Paper presented at 91st Annual Meeting of the American Association of Geographers, March 14–18, Chicago.*

Can, A. (1996). Weight matrices and spatial autocorrelation statistics using a topological vector data model. *International Journal of Geographical Information Systems* 10, 009–1017.

Can, A. (1998). GIS and spatial analysis of housing and mortgage markets. *Journal of housing Research* 9, 61–86.

Costanigro, M., Mittelhammer, R.C. and McCluskey, J.J. (2009), Estimating Class-specific Parametric Models under Class Uncertainty: Local Polynomial Regression Clustering in a

Hedonic Analysis of Wine Market. *Journal of Applied Econometrics, Early View*

Frizado, J., Smith, B.W., Carroll, M.C. and Reid, N. (2009), Impact of polygon geometry on the identification of economic clusters, *Letters in Spatial and Resource Sciences*, 2, 31 – 44

Garretsen, H. and Peeters,J. (2009), FDI and the Relevance of Spatial Linkages: Do Third-Country Effects Matter for Dutch FDI? *Review of World Economics, July 2009, v. 145, iss. 2, pp. 319-38*

Jaffe, A.B, Trajtenberg, M. and Henderson, R. (1993), Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, *The Quarterly Journal of Economics, vol. 108, No.3., 577 – 598*

Kaye-Blake, W., O'Connell, A. and Lamb, C. (2007), Potential Market Segments for Genetically Modified Food: Results from Cluster Analysis, *Agribusiness, Vol. 23 (4), 567 – 582*

Noev, N., (2005), Wine Quality and Regional Reputation: Hedonic Analysis of the Bulgarian Wine Market, *Eastern European Economics*, *Volume 43, Number 6 / November-December 2005, 5 – 30*

Porter, M. (2000), Location, Competition, and Economic Development: Local Clusters in a Global Economy. *Economic Development Quarterly, Vol.14 No.1,* 15-34.

Steiner, B., (2004), French Wines on the Decline? Econometric Evidence from Britain, *Journal of Agricultural Economics, Volume 55, Number 2, July 2004 , pp. 267-288(22)*

Stiler, J.F. (2007), *The House of Mondavi: The Rise and Fall of an American Wine Dynasty.* Penguin, New York, NY.

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46: 234–40.

Wallsten, S.J. (2001), An empirical test of geographic knowledge spillovers using geographic information systems and firm-level data. *Regional Science and Urban Economics 31, 571 – 599*

Wu, Y., and Hendrick, R. (2009). Horizontal and Vertical Tax Competition in Florida Local Governments. *Public Finance Review, 37,* 289

Table 4.1 Descriptive statistics of the non-binary variables for Washington and Calfornia

Descriptive Statistics

State		Price	Score	Case	Age
Washington (N = 79)	Mean	25.18	86.53	3052.28	2.81
	Min	10.65	78.00	110.00	2.00
	25 quartile	17.86	85.00	296.50	2.50
	Median	23.73	86.67	793.75	2.90
	75 quartile	29.75	88.39	1706.97	3.05
	Max	59.23	92.35	86321.46	4.17
	Std.	10.26	2.88	10153.72	0.52
California (N = 876)	Mean	34.88	85.40	4498.28	2.83
	Min	5.85	70.00	50.00	1.00
	25 quartile	18.00	83.42	450.00	2.43
	Median	25.58	85.91	975.30	2.93
	75 quartile	38.00	87.62	2764.06	3.09
	Max	1267.78	96.00	328333.33	5.50
	Std.	57.06	3.65	16438.23	0.59

* CPI adjusted to 2000

Table 4.2 Brief Descriptions and Abbreviations of Variables

Variables	Short Description	Binary/Non-binary
Score	Rating Score from Wine Spectator	Non-binary
Case	Number of Cases Produced	Non-binary
Age	Years of Aging Before Commercialization	Non-binary
Napa		
BayCentral		
Sonoma		
SouthCoast	Regions of Production in California	Binary
Carneros		
SierraFoothills		
Mendocino		
Columbia Valley		
Yakima Valley		
Walla Walla Valley	Regions of Production in Washington	Binary
Puget		

Table 4.3 Moran's I Tests for Washington State

Models	Variables	I	Sd (I)	Z	P-value
1 nearest neighbor	Price	0.582	0.14	4.239	0.000
	Score	0.209	0.139	1.593	0.056
2 nearest neighbors	Price	0.524	0.099	5.43	0.000
	Score	0.236	0.098	2.538	0.006
3 nearest neighbors	Price	0.526	0.082	6.603	0.000
	Score	0.287	0.081	3.71	0.000
4 nearest neighbors	Price	0.502	0.07	7.332	0.000
	Score	0.235	0.069	3.572	0.000
5 nearest neighbors	Price	0.484	0.062	8.007	0.000
	Score	0.275	0.062	4.677	0.000

Table 4.4 Moran's I Tests for California

Models	Variables	I	Sd (I)	Z	P-value
1 nearest neighbor	Price	0.454	0.042	10.737	0.000
	Score	0.31	0.042	7.347	0.000
5 nearest neighbors	Price	0.395	0.02	20.305	0.000
	Score	0.284	0.02	14.635	0.000
10 nearest neighbors	Price	0.385	0.014	27.824	0.000
	Score	0.272	0.014	19.65	0.000
20 nearest neighbors	Price	0.377	0.01	38.479	0.000
	Score	0.254	0.01	26.03	0.000
30 nearest neighbors	Price	0.363	0.008	45.611	0.000
	Score	0.24	0.008	30.208	0.000
35 nearest neighbors	Price	0.362	0.007	49.345	0.000
	Score	0.24	0.007	32.72	0.000
50 nearest neighbors	Price	0.342	0.006	56.096	0.000
	Score	0.225	0.006	37.073	0.000
60 nearest neighbors	Price	0.334	0.006	60.596	0.000
	Score	0.221	0.006	40.092	0.000

Table 4.5 Spatial Regression results for Washington State

Variables	Spatial Models		
	1 nearest neighbor	2 nearest neighbor	**3 nearest neighbor**
Intercept	-3.1014	-3.1773	**-3.3642**
	(0.001)	(0.001)	**(0.000)**
Score	0.0625	0.0585	**0.0591**
	(0.000)	(0.000)	**(0.000)**
case	-5.45E-06	-5.15E-06	**-5.66E-06**
	(0.056)	(0.065)	**(0.038)**
Age	0.1482	0.1591	**0.1538**
	(0.008)	(0.004)	**(0.004)**
Columbia Valley	-0.2406	-0.2217	**-0.2295**
	(0.023)	(0.032)	**(0.023)**
Yakima Valley	-0.1219	-0.1109	**-0.1193**
	(0.221)	(0.255)	**(0.212)**
Walla Walla Valley	0.1206	0.1043	**0.0564**
	(0.308)	(0.362)	**(0.629)**
Puget	-0.0007	-0.0183	**-0.0359**
	(0.994)	(0.846)	**(0.701)**
Rho	0.1572	0.2770	**0.3314**
	(0.085)	(0.013)	**(0.003)**
Wald test	2.9720	6.1290	**8.6570**
	(0.085)	(0.013)	**(0.003)**
LR test	2.9180	5.8330	**8.1190**
	(0.088)	(0.016)	**(0.004)**
LM test	1.9820	5.0480	**8.6620**
	(0.159)	(0.025)	**(0.003)**
AIC	20.8122	17.8973	**15.6105**

*P-values are in parenthesis

Table 4.6 Spatial Regression results for Washington State (continue)

Variables	4 nearest neighbor	5 nearest neighbor	No Spatial
Intercept	-3.4679	-3.6567	-3.0755
	(0.000)	(0.000)	(0.004)
Score	0.0603	0.0602	0.0676
	(0.000)	(0.000)	(0.000)
case	-5.49E-06	-5.09E-06	-5.85E-06
	(0.047)	(0.064)	(0.059)
Age	0.1482	0.1540	0.1561
	(0.006)	(0.004)	(0.011)
Columbia Valley	-0.2383	-0.2288	-0.2777
	(0.019)	(0.024)	(0.015)
Yakima Valley	-0.1254	-0.1208	-0.1351
	(0.194)	(0.208)	(0.209)
Walla Walla Valley	0.0373	0.0159	0.1840
	(0.760)	(0.897)	(0.132)
Puget	-0.0481	-0.0585	0.0108
	(0.615)	(0.540)	(0.916)
Rho	0.3395	0.3989	N/A
	(0.007)	(0.004)	
Wald test	7.2370	8.4290	N/A
	(0.007)	(0.004)	
LR test	6.8640	7.9440	N/A
	(0.009)	(0.005)	
LM test	7.5470	8.5740	N/A
	(0.006)	(0.003)	
AIC	16.8665	15.7856	19.7300

*P-values are in parenthesis

Table 4.7 Spatial Regression results for California

Variables	Spatial Models			
	1 nearest neighbor	5 nearest neighbors	10 nearest neighbors	20 nearest neighbors
Intercept	1.2966	1.2008	1.1998	1.1602
	(0.000)	(0.000)	(0.000)	(0.000)
Score	-0.0097	-0.0093	-0.0094	-0.0093
	(0.000)	(0.000)	(0.000)	(0.000)
case	5.13E-07	5.12E-07	5.01E-07	5.08E-07
	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.0153	-0.0149	-0.0149	-0.0147
	(0.000)	(0.000)	(0.000)	(0.000)
BayCentral	-0.0299	-0.0302	-0.0317	-0.0341
	(0.000)	(0.000)	(0.000)	(0.000)
Carneros	-0.0534	-0.0525	-0.0526	-0.0510
	(0.000)	(0.000)	(0.000)	(0.000)
Mendocino	-0.0162	-0.0166	-0.0167	-0.0178
	(0.0540)	(0.0460)	(0.045)	(0.031)
Napa	-0.0443	-0.0384	-0.0372	-0.0345
	(0.000)	(0.000)	(0.000)	(0.000)
SierraFoothills	-0.0068	-0.0081	-0.0089	-0.0094
	(0.3540)	(0.2650)	(0.219)	(0.195)
Sonoma	-0.0244	-0.0227	-0.0228	-0.0220
	(0.000)	(0.000)	(0.000)	(0.000)
SouthCoast	-0.0171	-0.0171	-0.0167	-0.0175
	(0.0010)	(0.0010)	(0.000)	(0.000)
Rho	0.1028	0.2366	0.2494	0.3110
	(0.000)	(0.000)	(0.000)	(0.000)
Wald test	19.4190	41.2870	39.4490	51.9940
	(0.000)	(0.000)	(0.000)	(0.000)
LR test	19.1960	40.3260	38.5550	50.4670
	(0.000)	(0.000)	(0.000)	(0.000)
LM test	13.9310	37.8090	44.4240	62.9270
	(0.000)	(0.000)	(0.000)	(0.000)
AIC	-3139.313	-3160.443	-3158.673	-3170.584

*P-values are in parenthesis

Table 4.8 Spatial Regression results for California (continue)

Variables	Spatial Models				
Variables	30 nearest neighbors	**35 nearest neighbors**	50 nearest neighbors	60 nearest neighbors	No Spatial
Intercept	1.1493	**1.1469**	1.14565	1.1461	1.3805
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.000)
Score	-0.0093	**-0.0093**	-0.0093	-0.0094	-0.0101
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.000)
case	5.09E-07	**5.07E-07**	5.08E-07	5.08E-07	5.40E-07
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.000)
Age	-0.0146	**-0.0145**	-0.0145	-0.0146	-0.0158
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.000)
BayCentral	-0.0341	**-0.0346**	-0.0360	-0.0360	-0.0321
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.000)
Carneros	-0.0513	**-0.0502**	-0.0504	-0.0502	-0.0589
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.064)
Mendocino	-0.0181	**-0.0174**	-0.0161	-0.0159	-0.0159
	(0.028)	**(0.034)**	(0.051)	(0.053)	(0.000)
Napa	-0.0341	**-0.0338**	-0.0335	-0.0330	-0.0503
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.352)
SierraFoothills	-0.0106	**-0.0119**	-0.0136	-0.0137	-0.0070
	(0.145)	**(0.099)**	(0.062)	(0.061)	(0.000)
Sonoma	-0.0218	**-0.0216**	-0.0217	-0.0214	-0.0269
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.001)
SouthCoast	-0.0183	**-0.0178**	-0.0180	-0.0178	-0.0181
	(0.000)	**(0.000)**	(0.001)	(0.001)	(0.000)
Rho	0.3345	**0.3406**	0.3585	0.3649	N/A
	(0.000)	**(0.000)**	(0.000)	(0.000)	
Wald test	54.1350	**55.2010**	54.3890	53.8100	N/A
	(0.000)	**(0.000)**	(0.000)	(0.000)	
LR test	52.4740	**53.4780**	52.7230	52.1660	N/A
	(0.000)	**(0.000)**	(0.000)	(0.000)	
LM test	69.3190	**71.5670**	71.3490	71.8850	N/A
	(0.000)	**(0.000)**	(0.000)	(0.000)	
AIC	-3172.591	**-3173.595**	-3172.84	-3172.283	-3124.117

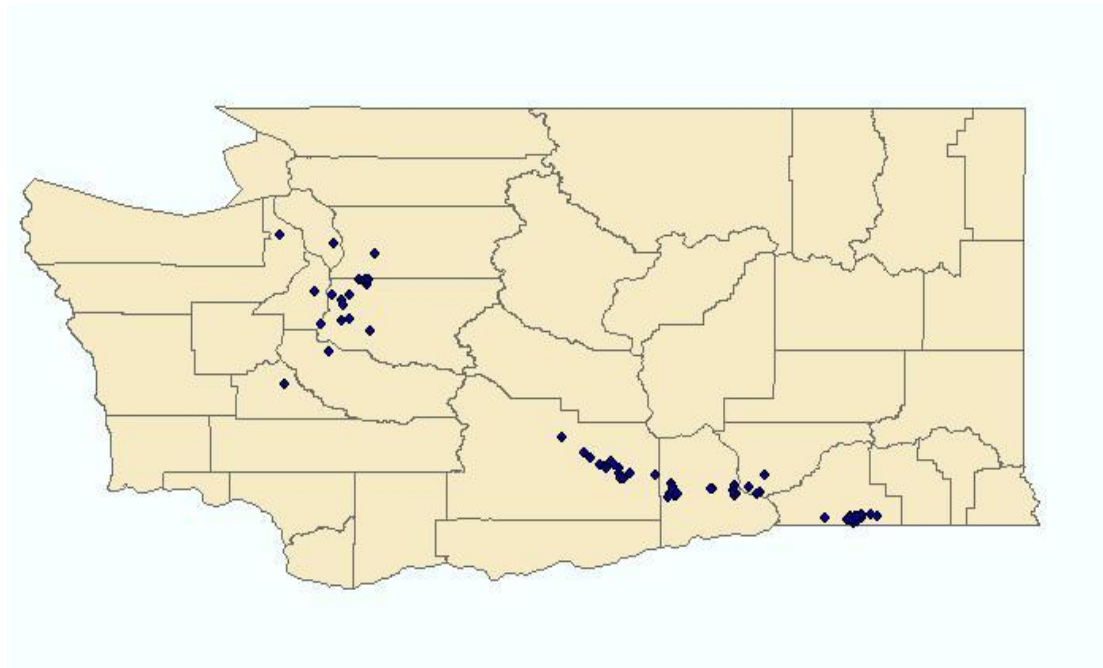*P-values are in parenthesis

Figure 0.1 Winery Distribution for Washington State

Figure 0.2 Winery Distribution for California State