Journal of Wine Economics (2022), **17**, 241–256
doi:10.1017/jwe.2022.46

ARTICLE

The effects of knowledge spillovers and vineyard proximity on winery clustering

Eric Stuen[1], Haifeng Liao[2] and Jon Miller[1]

[1]Department of Business, University of Idaho, 875 Perimeter Dr., MS 3161, Moscow, ID, 83844 and [2]Department of Geography and Geological Sciences, University of Idaho, 875 Perimeter Dr., MS 3021, Moscow, ID, 83843

Corresponding author: Eric Stuen, email: estuen@uidaho.edu

Abstract

We study the effect of proximity to other wineries on the formation of new wineries and how this effect depends on winemaking history in a location. Clustering is common in the wine industry, but it also depends on other factors, such as proximity to vineyards and high-reputation wineries. Using panel data with annual observations from 1994 to 2014 on 598 zip codes within Washington State, we estimate empirical models that control for proximity to wineries, proximity to vines, proximity to income, and the presence of star wineries. We find that the elasticity of the number of wineries with respect to proximity to wineries outside the zip code hinges on the length of local winemaking history. For locations with 11 or more winery years prior to our sample, the elasticity is at least 0.44. The presence of elite wineries is also found to have an effect, with about 0.5 additional wineries per year starting in a zip code per star winery. The effect of history suggests that policies to seed winery start-ups will help cluster formation, but only with a substantial critical mass of winemaking activity.

Keywords: clusters; knowledge network; knowledge spillovers; wine industry

JEL Classifications: O1; O18; O33; R11

I. Introduction

One of the central questions in economic geography is why industries tend to cluster. The factors of technological spillovers, proximity to intermediate inputs, size and composition of the local labor pool, and market access for final products were identified by Marshall (1920) and have been widely studied, spawning thousands of studies across the disciplines of economics, management, and geography (Hervas-Oliver et al., 2015). However, empirical assessment of the relative importance of these factors has been difficult to make for the simple reason that knowledge spillovers are hard to observe and quantify. In this study, we conduct an empirical assessment that accounts for many of these factors simultaneously.

Winemaking is an interesting industry in which to study these competing forces. It has been noted for the strong localization of its intermediate inputs, leading to certain regions having a strong natural geographic advantage over others. In a seminal article, Porter (1998) used wine to illustrate his theory of industrial clustering, noting that wineries depend on a host of localized inputs such as vineyards, general agricultural activity, wine tourism promotion, and specialized wine-making equipment. Ellison and Glaeser (1999) singled out the wine industry as highly dependent on local intermediate inputs and the second-most agglomerated industry in the United States. Yet, winemaking is also knowledge-intensive, and the application of scientific techniques has been increasing in recent decades (Jackson, 2014). Chemical analysis of wines is now standard, but there is an artistry to the process as well, from harvesting to crushing through blending to aging and bottling, all of which goes into the creation of wines of unique character. Such artistry has a strong tacit dimension in that it can only be communicated through hands-on demonstrations and hence locally. Local knowledge spillovers are therefore likely to be highly important in the wine industry (Galbreath, 2016).

Washington does not have a long history as a major wine-producing state but has rapidly become the second-largest wine-producing state in the United States. As recently as the 1980s, fewer than 30 wineries were operating in the state. The industry has grown substantially, and as of 2018 had 792 wineries in operation. These wineries have largely clustered into two areas. The largest stretches from Walla Walla County through Benton to Yakima County. In 2014, 41% of Washington's total winemaking firms were located in that cluster, employing an estimated 46% of the industry's workforce within the state. The second largest cluster is around the suburban town of Woodinville in King County, which accounted for an estimated 27.7% of the workforce. The remainder of Washington's wine industry is scattered across the rest of the state, with no county having more than 10 firms, except Chelan, with a small cluster of 15 firms. Figure 1 shows the distribution of wineries across zip codes in Washington State in 2014. In it, we see that the zip codes with the highest number of wineries tend to be surrounded by zip codes with an elevated number of wineries.

The remarkable thing about Washington's two major wine clusters is how different they are in terms of climate. King County (in which Seattle is located) is cloudy, damp, and relatively cold. Data on the Huglin climate index (Huglin, 1978), which is based on daily mean and maximum temperatures, indicates that Woodinville is unsuitable for most wine grape varieties in 71% of years. The other 29% of the time, it is warm enough for early-ripening varieties such as Riesling, Pinot Noir, and Chardonnay to reach maturity. In short, it is not a prime location for viticulture. On the other hand, Benton County is rated as temperate to warm 95% of the time, which provides an ideal growing climate, similar to parts of Spain and Italy.

Why does a relatively unsuitable place for grape production like King County have one of the state's major winery clusters? Its history illustrates the path dependence in winery clustering. After the end of prohibition, Woodinville was a center of cheap wine production based on Concord grapes that could handle King County's damp climate. In 1964, Associated Vintners was founded, the state's first winery to emulate old-world wines made with *vitis vinifera* grapes.[1] It was started by faculty from the nearby University of Washington—an infusion of human capital. The old Concord grape wineries were consolidated to form a new winery, Chateau Ste. Michelle,

[1]The species *vitis vinifera* includes all varieties associated with traditional European wine production, such as Cabernet Sauvignon, Merlot, Grenache, Syrah, etc.

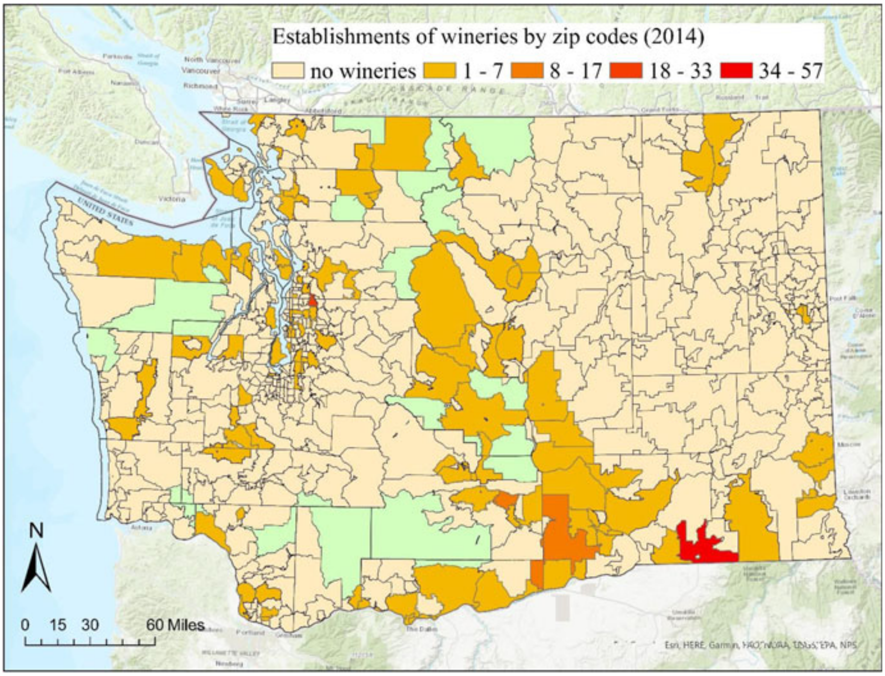

Figure 1. Distribution of wineries by zip codes, 2014.
Source: County Business Patterns Data, U.S. Census Bureau and ESRI ArcGIS Map 2.6 Software. Copyright the Authors.

which also turned to premium wines (MacNeil, 2001). Associated Vintners was later renamed Columbia Winery. These two wineries are the oldest still operating in Washington. The advantages that Woodinville provided, such as access to winemaking equipment, winemaking knowledge, and proximity to the Seattle-Tacoma metropolitan market, evidently outweighed the disadvantage of being 200–300 miles away (and over a mountain range) from most vineyards that are ideal for *vitis vinifera*. Figure 2 shows the overlap of counties and American Viticulture Areas (AVA) in Washington State. The AVA in the eastern half of the state is subdivided into ten distinctive appellations due to the varied geography of the region.

For our analysis, we constructed measures of the overall proximity of zip codes to factors that are likely to affect the location of new wineries. Proximity to wineries in other zip codes is our proxy for proximity to winemaking knowledge. Another key variable measures proximity to vineyards. Indexes for each of these were constructed using distances from each zip code as weights.

It is difficult to quantify the amount of production knowledge held locally. We think that production knowledge in the wine industry is proportional to the number of wineries because most U.S. wineries employ a small team of winemakers.[2]

[2]Through conversations with wine-industry professionals, the authors have learned that it is typical for small- and medium-sized wineries to employ just one to three winemakers.

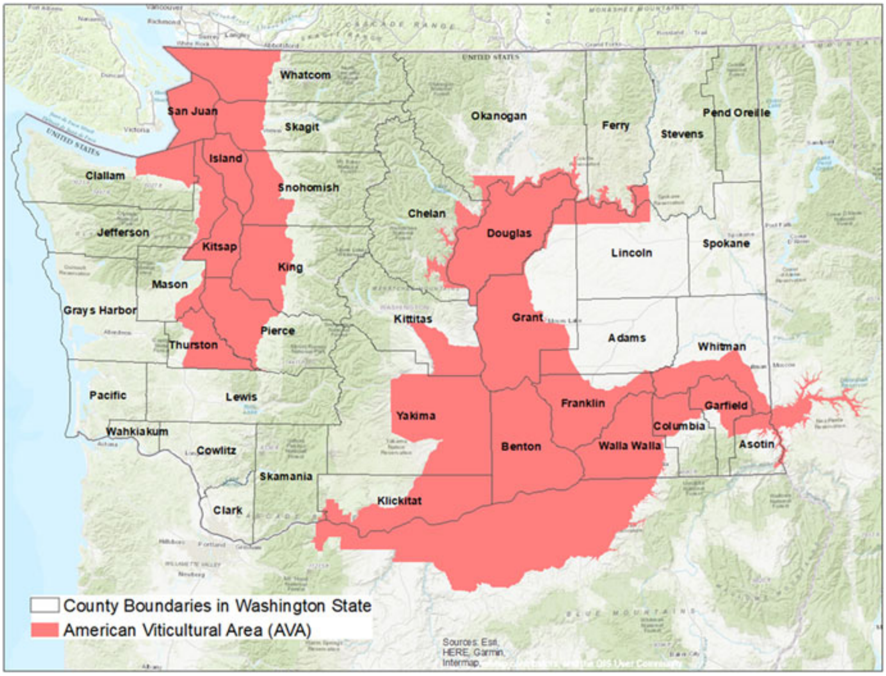

Figure 2. Locations of AVA in the state of Washington.
Source: U.S. Alcohol and Tobacco Tax and Trade Bureau, and ESRI ArcGIS Map 2.6 Software. Copyright the Authors.

Winemakers are responsible for creating wines of distinctive quality, monitoring each part of the winemaking process. They frequently test the wines both through chemical analyses and taste tests. While some high-end wineries hire expert consultants, that is, "flying winemakers," to give input on the process, most of the winemaking process is managed by local winemakers. Therefore, the amount of local knowledge must be proportional to the number of wineries.

This is not to imply that we think the winemaking knowledge of each winery is equal. Talent varies, both that which is developed within the region and from experienced winemakers who have come from elsewhere. To further control winemaking talent, we identify "best" wineries based on the tail end of the distribution of wine critic ratings. We include a measure of the number of best wineries per zip code in our analyses.

We estimate regressions of the change in the number of wineries on proximity indexes for other wineries, vineyards, and income, as well as the number of "best" wineries. Proximity to wineries has a positive effect on winery formation, but only in locations with a significant prior history of winemaking. The presence of elite wineries is also seen to have a positive effect on the number of wineries. These effects are also found in negative binomial and zero-inflated negative binomial (ZINB) models using the level of wineries as the dependent variable.

The effect of proximity to other wineries hinges on the presence of early wineries: for zip codes with at least 11 years of winery operation prior to the sample timeframe, the elasticity of the number of wineries with respect to proximity to wineries outside the zip code is at least 0.44, and is increasing in the number of prior years.[3]

The effect of proximity to other wineries could be due to several mechanisms, from which we are not able to differentiate. The first is knowledge transfer through official channels such as inter-winery cooperation, consulting, and trade fairs. The second is knowledge "spilled" through employee turnover, a benefit of deeper labor-market pools. The third is the presence of other inputs, such as specialized equipment suppliers. So, while we cannot isolate the effect of knowledge spillovers, our results indicate a general momentum in winery clusters that could be attributed to them. Our statistical evidence, which is based on data aggregated to zip code, AVA, and county levels, complements the micro evidence from other studies that survey winemakers directly, such as those cited later.

The rest of the paper is organized as follows. Section II discusses related literature. In Section III, we describe our empirical methodology. Section IV describes our data, and Section V our results. We offer concluding remarks in Section VI.

II. Related literature

Our study contributes empirical evidence on knowledge spillovers in winemaking, but it also contributes to a broader literature on how knowledge spillovers contribute to localization economies. It most directly adds to the handful of studies that have documented knowledge transmission within winery clusters using detailed survey data. Prominent among those studies are Giuliani and Bell (2005), Morrison and Rabellotti (2005), Giuliani (2007), Farias and Tatsch (2014), Maghssudipour, Lazzeretti, and Capone (2020), and Choi and Gu (2020).

Giuliani and Bell (2005) surveyed 32 firms in Chile and analyzed the characteristics of the social network between them, emphasizing formally acknowledged knowledge-sharing ties. They also highlighted the role of "flying winemakers," consultants who work for several wineries and are not bound by location. Similarly, Morrison and Rabellotti (2005) surveyed 26 wineries in an Italian cluster and mapped social proximity between the firms, finding that some firms practice a mutual exchange of knowledge while others are isolated. Giuliani (2007) analyzed survey data from 105 wineries in 3 clusters (2 in Italy, 1 in Chile), and mapped the self-reported linkages between firms in terms of both winemaking and business knowledge. She found that firms with more human capital tend to be more linked, but only with firms with a similar knowledge base due to the need for reciprocal knowledge sharing. Farias and Tatsch (2014) conducted a similar style of study as the previously mentioned of 20 wineries in a Brazilian winery cluster and found similar results.

Recent studies by Maghssudipour, Lazzeretti, and Capone (2020) and Choi and Gu (2020) focus on the multiplicity of knowledge networks that wineries are connected to. Maghssudipour, Lazzeretti, and Capone distinguished between economic and social ties in studying the Montefalco cluster in Italy, finding that economic ties are more important to knowledge transfer. Choi and Gu focused on the linkages

[3]The "prior years" variable is aggregated over all wineries in the zip code, that is, one winery operating for ten years is equivalent to two wineries operating for five years.

between universities, government agencies, and private companies in innovation in China's nascent wine industry. They find that winemaking knowledge first clustered around China's coastal cities, then spread inland during the period 2007–2016.

Another impetus for starting a winery in an established wine region is to take advantage of the collective reputation that the region's wineries enjoy. Collective reputation has been previously studied by numerous authors, see Tirole (1996) for a seminal theoretical treatment of it, and Costanigro, Bond, and McCluskey (2012) on the simultaneous effects of collective and private reputation. Yang, McCluskey, and Brady (2012) found evidence that clustered wineries are positively correlated in terms of prices and wine ratings, and they used hedonic spatial lag models to estimate the effect of proximity, weighted by quality, on wine prices. This provides detailed microeconomic evidence that proximity to high-quality wineries is a determinant of wine quality.

The Marshallian externality of intra-industry knowledge spillovers includes all knowledge transfer, not just those through formal channels, and studies that rely on survey data may miss the more indirect channels of knowledge transfer. In the wine industry, other channels include churn in the market for skilled labor (experienced winemakers leaving one firm for another, or starting their own), unofficial personal contacts, and trade conferences. The empirical method applied in our paper captures knowledge spillovers indirectly, inferring that the effect of proximity to other wineries must be due to knowledge transfer, once other factors are controlled for. A downside to indirect methods such as ours is that they may conflate knowledge spillovers with unobserved externalities driving location, a point made by Breschi and Lissoni (2001). While we try to control for other factors driving localization, we acknowledge that there is no perfect way to do so.

As mentioned earlier, the idea that firms cluster near other firms in the same industry due to the "Marshallian externalities" of proximity to intermediate input suppliers, local labor pool, and intra-industry technological spillovers was first described by Marshall (1920), later reemphasized by Krugman (1991). Arrow (1962) outlined the economics of innovation, noting that an inherent disincentive for R&D activity is the potential for knowledge to spill to rival firms, especially through labor-market churn. That such intra-industry technological spillovers generate increasing returns to scale at the industry level later motivating the development of models of endogenous economic growth (e.g., Romer 1986, 1990).

Rosenthal and Strange (2004) have a detailed discussion of the literature regarding empirical studies of the relative importance of the Marshallian external factors for industry concentration. Our study adds to this literature. Jaffe (1989) specified university and private R&D as inputs into state-level knowledge production functions. Jaffe, Trajtenberg, and Henderson (1993) found that patents are more likely to be cited by other patents within the same state and metropolitan area. Audretsch and Feldman (1996) found that the more knowledge-oriented an industry, the more geographically clustered its innovative activity, which is indicative of the presence of local knowledge spillovers. Rigby and Essletzbichler (2002) examined national plant-level data and created direct measures of the external factors, finding positive effects of the concentration of supply linkages, productivity growth of suppliers (a proxy for knowledge), and labor-pool mix. Holmes (1999) looks at a single industry (textiles) and finds that firms located in areas with high concentrations of other firms in the industry are more vertically disintegrated, that is, they purchase inputs from supplying firms more often than they vertically integrate the supply. Evidence shows that proximity to inputs is an important positive localization externality. More recently, Bloom, Schankerman, and Van Reenen (2013) argued that technology spillovers are offset by the effects of competitor firms carrying out R&D and taking market share from product market rivals. They estimated the technology spillover effect to be more than twice as high as the market share effect at the aggregate level. Another detailed empirical analysis was done by Greenstone, Hornbeck, and Moretti (2010), who found a large productivity effect on incumbent firms from the quasi-random placement of new plants nearby.

III. Empirical methodology

We test the relationship between the number of winemaking establishments (wineries) per zip code and measures of the strength of externalities operating in that location. To do so, we constructed a panel dataset at the zip code year level, covering 598 zip codes in Washington State and the years from 1994 to 2014. Our zip code coverage was limited to those for which we had access to data on the distances between them, and our panel only includes 83% of the 719 zip codes within Washington State. However, only 179 of the zip codes ever had one or more wineries, and almost all of these are included in the 598.

Analysis at the zip code level allows us to account for the effects of proximity at a highly granular level. The size of zip codes varies with population density, but for Washington, most zip codes are less than ten miles wide. This helps to limit the measurement error in our study: the variation in distances between zip codes (measured) is much larger than the variation in distances within zip codes (unmeasured and assumed to be zero).

Each proximity variable is constructed in the following manner:

$$index_{it} = \sum_{j=1}^{J} \frac{x_{jt}}{d_{ij}},$$

where i indicates zip code, t indicates year, x indicates a variable that is dispersed over a geographic unit j, and d is the distance between j and zip code i.

Knowledge spillovers in this context refer to the use of winemaking knowledge by other firms than those that generated it. It is usually the case that each winery has a team of winemakers. Presumably, each winemaker has idiosyncratic knowledge regarding the winemaking craft. Therefore, it is reasonable to presume that the quantity of winemaking knowledge within a zip code is an increasing function of the number of wineries near that zip code in a region. Our distance-weighted measure of winemaking knowledge is therefore:

$$WI_{it} = \sum_{k=1}^{K} \frac{w_{kt}}{d_{ik}},$$

where w_{kt} is the number of wineries in zip code $k \neq i$ in year t. This index applies distance weights linearly.

Locations close to vineyards provide a natural advantage for wineries, which can better monitor and protect the quality of grapes, as well as save on transport costs. We construct an index of proximity to vineyards in the form:

$$VI_{it} = \sum_{j=1}^{J} \frac{vines_{jt}}{d_{ij}}$$

$Vines_{jt}$ is the number of grapevines cultivated in AVA j in year t.[4] The factor d_{ij} is the Euclidean distance d between zip code i and AVA j. Ideally, this measure would be constructed using distances from each winery to each acre of wine grapes. As constructed here, however, we use the centroid distances from the approximate center of each zip code to the center of each AVA. Since Washington's AVA vary in size and shape, this measure has an upward measurement error for some zip codes and a downward measurement error for others.

Another factor we consider is the distance to customers, reflecting the influence of transportation costs. We constructed an income index using the income-weighted population sizes of counties in Washington.

$$II_{it} = \sum_{c=1}^{C} \frac{inc_{ct}}{d_{ic}}$$

Here, *inc* indicates aggregate household income in county c and year t. Distances are again measured using centroid distances, but this time between zip codes and counties.

To measure the depth of winemaking knowledge based on winemaking history in an area, we define the variable $Yearsprior_i$ as the number of years that early Washington wineries had collectively operated in zip code i prior to 1994. We define early Washington wineries as those in operation prior to 1985 that were still in business in 2017. In our models, we include the interaction between WI_{it} and $Yearsprior_i$, to allow the effect of proximity to wineries to depend on winemaking history.

As mentioned previously, a reputation for excellence among the wineries in a particular location may also attract new wineries. That is, critically acclaimed wineries in an area may raise the collective reputation of a winemaking region and attract more winemakers. This collective reputation effect is distinct from knowledge spillovers in that even wineries with no interaction with other winemakers can free-ride on their common appellation label. To examine the effect of the reputation of wineries on the growth of wineries, we focus on star wineries that have been highly critically acclaimed. We categorize a winery as a "best Washington winery" if it has produced at least one wine that is listed in Wine Searcher's "Best Washington Wines" list.[5]

[4] AVA are officially designated wine-growing regions for which regulations on wine labeling apply. The AVA is the smallest geographic unit for which we could obtain data on the vineyard area.

[5] www.wine-searcher.com

Details of the construction of this variable are given in the data section, but the definition of the variable is given as:

$$bestwa_{it} \equiv number\ of\ best\ wineries\ located\ in\ zip\ code\ i\ by\ year\ t$$

Our empirical approach starts with regression models estimating the effects of the variables on the change in the number of wineries per zip code. This model takes the form:

$$\Delta w_{it} = \beta_0 + \beta_1 VI_{it} + \beta_2 II_{it} + \beta_3 WI_{it} + \beta_4 WI_{it} * YP_i + \beta_5 bestwa_{it} + \boldsymbol{\alpha}_t + \epsilon_{it}, \quad (1)$$

where w is the number of wineries, *VI, II,* and *WI* are the proximity indexes for vineyards, income, and other wineries, *YP* is the history variable *yearsprior,* and alpha is a set of year indicators.

We also estimate the model in levels to examine the co-trending behavior of the number of wineries and the explanatory variables. Due to the high number of zip code year observations with zero wineries, we turn to negative binomial regression models. Negative binomial models are suitable for this data because the dependent variable is a discrete count with many zero-valued observations.

In general, negative binomial models are specified by

$$f(w_{it}|\boldsymbol{x_{it}}, u_{it}) = \frac{e^{-\lambda_{it}u_{it}}(\lambda_{it}u_{it})^{w_{it}}}{w_{it}!}, \quad (2)$$

where $\boldsymbol{x_{it}}$ is a vector that includes all of the right-hand side variables in Equation (1) and

$$\boldsymbol{x_{it}}'\boldsymbol{\beta} + \epsilon_{it} = \ln \lambda_{it} + \ln u_{it}. \quad (3)$$

Here, the parameter λ_{it} is the expected number of wineries per zip code and year, and u_{it} is a random effect that is usually assumed to have a gamma distribution.

IV. Data

Our primary source of data on the number and size of wineries is the *County Business Patterns* (*CBP*) survey, a publicly available dataset from the U.S. Census Department. The *CBP* contains data on the number of wine-producing establishments (plants) at the zip code level for the years 1994–2014. It also contains the number of wine establishments in each of many employment size categories, from which we can construct a rough measure of winemaking employment per zip code.

To construct our *Yearsprior* variable, we gathered information on early Washington wineries from winery licenses issued by the Washington State Liquor and Cannabis Board, and trade publications. To our knowledge, there are only 23 wineries that meet our criteria of having been established prior to 1985 and surviving to the present. The relative youth of the Washington wine industry means that early in its history, winemaking knowledge was scarce and sparsely distributed. According to our data, among Washington zip codes that have ever had a winery, the mean *Yearsprior* is 4.6. However, two zip codes have more than ten times that: 98072, Woodinville, and 99350, Prosser.[6]

For the variable *vines_index*, our index of proximity to grapes, we used data from the USDA National Agricultural Statistics Service (2011). This report provided data on the quantity of vines in cultivation in each AVA for the years 1991, 1995, 2000, 2005, 2007, 2009, and 2010. For years between those dates, we estimated the quantity using linear interpolation. For the four years of our panel post-2010, we assumed no growth, and so used the 2010 values throughout. This assumption is reasonable because growth in the number of vines in cultivation dropped off after 2007 with the onset of the Global Financial Crisis. The distance data were constructed using a GIS to calculate distances between the approximate center of each zip code and AVA. Similarly, the variable inc_index_{it} was constructed using data on aggregate personal income per county from the Bureau of Economic Analysis and weighted with distances between zip codes and counties. It was deflated and converted to real income using the consumer price index for all urban consumers (CPI-U) from the Bureau of Labor Statistics. For the full sample of 12,558 observations for which there was data on the vineyard proximity index, Table 1 contains a summary of all the data used in this analysis.

To construct the *bestwa* (best Washington wineries) variable, we turned to Wine Searcher's list of "Best Washington Wines." For each wine, Wine Searcher collected critic scores from dozens of leading publications. While there is a lot of noise in individual critics' scores (as shown in Stuen, Miller, and Stone (2015)), the averages provided by Wine Searcher are much more reliable. Wine Searcher's Best Washington Wines list includes 25 wines that average 93 or above on the common 100-point rating scale. The determination of the "best" wines and wineries is, like judging art, inherently subjective, and we know that many excellent wineries are left off this list. In defining "best" too narrowly, we risk limiting the explanatory power of the variable, but we do not find that to be the case in the empirical results.

Those 25 wines are produced by just 11 wineries, operating in just 5 zip codes. These wineries and their zip codes of operation are listed in Table 2.[7]

Two of these are located in the Woodinville cluster (which includes nearby Snohomish), while seven are in Walla Walla. Three are co-categorized as "early wineries," having begun operations prior to 1985 (these being Quilceda, Leonetti, and Chateau Ste. Michelle).

V. Results

Table 3 shows the regression results for the models with the change in the number of wineries as the dependent variable, estimating the model in Equation (1). The model in Column (1) excludes the *bestwa* variable. We see from that model run that the

[6]Early wineries in Woodinville include Columbia and Chateau Ste. Michelle. Early wineries in Prosser include Hinzerling, Yakima River, Hogue, Chinook Yakima Valley, and Pontin Del Roza.

[7]Zip codes of operation were determined by addresses listed on winery websites since the Census *CBP* data that we use to construct winery counts is anonymized.

https://doi.org/10.1017/jwe.2022.46 Published online by Cambridge University Press

Table 1. Descriptive statistics (Sample size: N = 12,558)

Panel dimensions: Years 1994–2014, 598 Washington zip codes					
Variable	Source	Non-zero observation %	Mean	Minimum	Maximum
Wineries	*CBP* Survey, U.S. Census Bureau	9.52	0.21	0	60
Vineyard proximity index	NASS vines/AVA data; GIS	100	167.48	20.23	2906.98
Income proximity index	BEA; GIS	100	12.06	2.55	248.78
Proximity to wineries, linear distance weights	*CBP*; GIS	100	0.33	0.03	2.92
Best Washington wineries	Wine Searcher	0.54	0.009	0	7
Winery-years prior to 1994	Authors' investigations	2.84	0.58	0	68

Table 2. Best Washington wineries (based on average critic ratings)

Winery	First year rated among best	Winery city	Winery zip code
Quilceda Creek	1990	Snohomish	98290
Leonetti Cellar	1998	Walla Walla	99362
Chateau Ste. Michelle and Dr. Loosen[a]	2000	Woodinville	98072
Andrew Will	2001	Vashon	98070
Cayuse	2002	Walla Walla	99362
Corliss	2004	Walla Walla	99362
K Vintners	2006	Walla Walla	99362
No Girls	2008	Walla Walla	99362
Horsepower	2011	Walla Walla	99362
Upchurch	2014	Benton City	99320
Hors Categorie	2014	Walla Walla	99362

Note: [a]A collaboration between the Chateau Ste. Michelle winery and winemaker Ernst Loosen, under the Eroica brand label.

interaction of *yearsprior* with the winery proximity index has a positive and statistically significant coefficient estimate of 0.029. Without the interaction, the winery proximity index is not significant, but for zip codes with an early history of winemaking, the effect is economically significant. The elasticity of wineries with respect to the winery proximity index is 0.04 for every winery year of winemaking history. For the 2% of zip codes with *yearsprior* values of 11 to 68, the winery proximity elasticity is therefore 0.44 to 2.72. This shows that in locations with early wineries, as the broader winemaking cluster grows, the heart of the cluster grows faster.

We add *bestwa* in model (2) to examine how the presence of winemaking stars influences winery growth. Its effect is estimated at 0.534 and has a straightforward interpretation: for every "best winery" in the zip code, the total number of wineries increases by 0.53 per year. Therefore, a zip with the four best wineries should see about two additional wineries per year, beyond the usual trend growth and growth due to other factors. After the inclusion of the *bestwa* variable, the coefficient on the interaction of the winery proximity index and *yearsprior* is diminished by 28% and is only marginally statistically significant, with a p-value above 0.05 but below 0.1. This signifies that some of the effect of the interaction variable was due to omission of winemaking reputation within the cluster as an independent factor.

In Columns (3) and (4) of Table 3, we examined whether the results were sensitive to the portion of the sample used to estimate them and found that they were. It is plausible that locations that have never had a winery are qualitatively different than those that have had one or a few, and that locations with many wineries are also qualitatively distinct. The sample for the model in Column (3) only includes zip codes that averaged at least one winery per year from 1994–2014, that is, the established zip codes. In Column (4), the sample only includes zip codes that averaged between zero and one winery (typically those that started with zero wineries and gained one or two during the period). The estimated effects in the established sample are slightly larger than in the full sample, which should be expected. The model has hardly any explanatory power in Column (4), with the interaction term having a slight negative effect, showing that winemaking history and prestige were largely irrelevant to why wineries were established in those zip codes.

Next, we turn to models of the total count of wineries per zip code regressed on the same explanatory variables as in Table 3. This essentially documents the co-evolution of multiple stock variables measuring the size of the wine industry in each location: wineries per zip code, the winery proximity index (size of the Washington wine industry, discounted for distance), and Best Washington Wineries, measuring the size of the winemaking elite. Since we include year effects in all specifications (as in Table 3), the estimated marginal effects are effects on local growth beyond the overall trend for the state.

Columns (1) and (2) of Table 4 show the zero and two-year lags of the negative binomial model estimation. As with the prior models, the interaction term and the *bestwa* variable have positive effects and are statistically significant. Here we also find the vineyard proximity index to be statistically significant. The magnitude of the winery proximity index effect is close to that which was estimated in model (2) of Table 3. Since negative binomial coefficients are interpreted as semi-elasticities, we look at 1% changes over the mean of each explanatory variable to see the elasticities in Column (1). The elasticity of wineries with respect to the winery proximity index is therefore 0.042 for each year of *yearsprior*. The elasticity with respect to the vineyard proximity index is 0.34. With respect to *bestwa*, it is 0.014. Column (2) shows the model estimated with two-year lags of the explanatory variables. The

Table 3. OLS regressions of the change in wineries per zip code

Sample	Full (1)	Full (2)	≥1 winery per year, average (3)	≥0 winery per year, and <1, average (4)
Vineyard proximity index	0.00010 (0.00007)	0.00008 (0.00005)	0.00014 (0.00016)	0.00002 (0.00003)
Winery proximity index	−0.019 (0.024)	−0.002 (0.014)	−0.148 (0.292)	−0.00090 (0.030)
Winery proximity index* *yearsprior*	0.029** (0.014)	0.021* (0.011)	0.025* (0.013)	−0.005** (0.002)
Income proximity index	−0.00004 (0.00020)	−0.00002 (0.00016)	0.003 (0.023)	0.00006 (0.00058)
Best Washington Wineries		0.534*** (0.047)	0.542*** (0.033)	−0.003 (0.014)
Intercept	0.00214 (0.02147)	−0.012 (0.026)	−0.009 (0.191)	0.033 (0.058)
Observations	11,960	11,960	480	2,860
R-squared	0.04809	0.130	0.206	0.013

Notes: Robust standard errors in parentheses. All regressions include year effects. ***p < 0.01, **p < 0.05, *p < 0.1

Table 4. Negative binomial and ZINB regressions of number of wineries on proximity indexes

Model	Negative binomial (1)	Negative binomial (2)	ZINB (3)
Vineyard proximity index	0.002*** (0.000)		0.002*** (0.000)
Winery proximity index	0.330 (0.314)		−0.527 (0.336)
Winery proximity index* *yearsprior*	0.132*** (0.036)		0.069*** (0.018)
Income proximity index	−0.008 (0.009)		0.000 (0.018)
Best Washington Wineries	1.609*** (0.404)		0.918*** (0.166)
Vineyard proximity index (2 year lag)		0.002*** (0.000)	
Winery proximity index (2 year lag)		0.363 (0.425)	
Winery proximity index* *yearsprior* (2 year lag)		0.157*** (0.044)	
Income proximity index (2 year lag)		−0.009 (0.009)	
Best Washington Wineries (2 year lag)		1.834*** (0.320)	
Intercept	−2.179*** (0.292)	−2.154*** (0.287)	−0.761 (0.463)
Observations	12,558	11,362	12,558
Pseudo R-sq	0.178	0.169	

Notes: Robust standard errors in parentheses. All regressions include year effects. ***p < 0.01, **p < 0.05, *p < 0.1

estimates are very similar and a bit stronger, confirming that the co-evolution of these measures is not sensitive to the lag structure.

In light of the fact that the count of wineries was zero for 90.5% of the zip code year observations in our sample, we also estimated a ZINB regression model in Column (3). This examines whether the variables have significant effects above and beyond the "hurdle" of having one or more wineries. The results are similar to the negative-binomial model, but the coefficients on the interaction variable and *bestwa* are truncated. The direct interpretation of the interaction coefficient from model (3) is that a 1% increase in the winery proximity index over its mean increases the number of wineries by 0.022% per early winery year, about half the effect estimated in Column (1). Overall, we can conclude that the variables explain both the presence and magnitude of wineries in each zip code.

Additional results are reported in an unpublished appendix, available upon request, in which robustness checks on zip code fixed-effects models, lagged models, and regressions of winery employment were examined. A list of the early wineries used to construct the *yearsprior* variable is also included.

https://doi.org/10.1017/jwe.2022.46 Published online by Cambridge University Press

VI. Concluding remarks

This study examines the spatial distribution of winemaking establishments, using anonymized data on the number of wineries in Washington State. It examines several critical factors: proximity to the primary raw material (grapes), proximity to consumer income, proximity to other wineries, winemaking history, and the presence of elite wineries. We find evidence that the effect of proximity to other wineries hinges on the local history of winemaking. For zip codes with a substantial history of winemaking (at least 11 winery years prior to 1994), the elasticity of the number of wineries with respect to proximity to other wineries in the region is at least 0.44. On the other hand, we find no such proximity effect for zip codes with no history of winemaking.

The finding that winery clustering depends on the length of winemaking history in a location has interesting policy implications. It suggests that winery clustering is not just the result of pre-existing geographic features such as the distribution of land suitable for viticulture and the distribution of the population. Historical accidents, such as chance meetings between winemakers, the affinity of experienced winemakers for particular locations, or the availability of other winemaking inputs (equipment adapted from other industries, technical education programs), may seed new clusters. Governments may therefore be able to encourage cluster formation through subsidies for new firms and education. However, efforts to encourage cluster formation in already-established wine regions, outside of existing clusters, face a substantial hurdle, as most economic incentives favor locating in existing clusters.

We also find evidence that the presence of star wineries has an effect on growth in the number of wineries. This effect could arguably go both ways, as elite winemakers could be attracted to growing winery clusters. We leave it to the reader to decide whether elite wineries are more of a cause or effect of the growth of a winery cluster.

Acknowledgments. The authors thank Olivier Gergaud, an anonymous reviewer, and participants in sessions at the 2018 AAWE conference in Ithaca, NY, and the 2017 WEAI conference in San Diego, CA, for helpful comments.

References

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In R. Nelson (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, 609–634. Princeton: Princeton University Press.

Audretsch, D., and Feldman, M. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86(3), 630–640.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4), 1347–1393.

Breschi, S., and Lissoni, F. (2001). Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change*, 10(4), 975–1005.

Choi, H., and Gu, C. (2020). Geospatial response for innovation in the wine industry: Knowledge creation through institutional mobility in China. *Agronomy*, 10(4), 495.

Costanigro, M., Bond, C., and McCluskey, J. (2012). Reputation leaders, quality laggards: Incentive structure in markets with both private and collective reputations. *Journal of Agricultural Economics*, 63(2), 245–264.

Ellison, G., and Glaeser, E. L. (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review*, 89(2), 311–316.

Farias, C., and Tatsch, A. (2014). The Brazilian wine industry: A case study on geographical proximity and innovation dynamics. *Revista de Economia e Sociologia Rural*, 52, 515–532.

Galbreath, J. (2016). Exploratory study of climate change innovations in wine regions in Australia. *Regional Studies*, 50(11), 1903–1918.
Giuliani, E. (2007). The selective nature of knowledge networks in clusters: Evidence from the wine industry. *Journal of Economic Geography*, 7(2), 139–168.
Giuliani, E., and Bell, M. (2005). The micro-determinants of meso-level learning and innovation: Evidence from a Chilean wine cluster. *Research Policy*, 34(1), 47–68.
Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3), 536–598.
Hervas-Oliver, J.-L., Gonzalez, G., Caja, P., and Sempere-Ripoll, F. (2015). Clusters and industrial districts: Where is the literature going? Identifying emerging sub-fields of research. *European Planning Studies*, 23(9), 1827–1872.
Holmes, T. J. (1999). Localization of industry and vertical disintegration. *Review of Economics and Statistics*, 81(2), 314–325.
Huglin, P. (1978). Nouveau mode d'évaluation des possibilités héliothermique d'un milieu viti-cole. *C. R. Académie d'Agriculture* (*Acad. Agric.*), 64, 1117–1126.
Jackson, R. (2014). *Wine Science: Principles and Applications*. London: Academic Press.
Jaffe, A. (1989). Real effects of academic research. *American Economic Review*, 79(5), 957–970.
Jaffe, A., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, 108(3), 577–598.
Krugman, P. (1991). *Geography and Trade*. Cambridge: MIT Press.
MacNeil, K. (2001). *The Wine Bible*. New York: Workman Publishing.
Maghssudipour, A., Lazzeretti, L., and Capone, F. (2020). The role of multiple ties in knowledge networks: Complementarity in the Montefalco wine cluster. *Industrial Marketing Management*, 90, 667–678.
Marshall, A. (1920). *Principles of Economics*. London: MacMillan.
Morrison, A., and Rabellotti, R. (2005). *Knowledge and Information Networks: Evidence from an Italian Wine Local System*. Milan, Italy: Università commerciale Luigi Bocconi.
Porter, M. E. (1998). Clusters and the new economics of competition. *Harvard Business Review*, 76(6), 77–90.
Rigby, D., and Essletzbichler, J. (2002). Agglomeration economies and productivity differences in U.S. cities. *Journal of Economic Geography*, 2(4), 407–432.
Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5), 1002–1037.
Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), S71–S102.
Rosenthal, S., and Strange, W. (2004). Evidence on the nature and sources of agglomeration economies. In V. Henderson and J. Thisse (eds.), *Handbook of regional and urban economics*, 4th ed., 2119–2171. North Holland: Elsevier.
Stuen, E., Miller, J., and Stone, R. (2015). An analysis of wine critic consensus: A study of Washington and California wines. *Journal of Wine Economics*, 10(1), 47–61.
Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, 63(1), 1–22.
USDA National Agricultural Statistics Service, Washington Field Office (2011). *Washington Vineyard Acreage Report 2011*. Available at https://www.nass.usda.gov/Statistics_by_State/Washington/Publications/Fruit/GrapesFinalwithCounty.pdf.
Yang, N., McCluskey, J., and Brady, M. (2012). The value of good neighbors: A spatial analysis of the California and Washington State wine industries. *Land Economics*, 88(4), 674–684.

Cite this article: Stuen, E., Liao, H., and Miller, J. (2022). The effects of knowledge spillovers and vineyard proximity on winery clustering. *Journal of Wine Economics*, 17(3), 241–256. https://doi.org/10.1017/jwe.2022.46