Geographical Indications and Welfare: Evidence from the US Wine Market

Raj Chandra, GianCarlo Moschini, Gabriel E. Lade

Working Paper 21-WP 628
Updated
August 2023

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Raj Chandra, currently at Amazon, worked on this paper as part of her dissertation while a PhD student at Iowa State University.

GianCarlo Moschini is Professor, Department of Economics, Iowa State University, Ames, Iowa 50014. E-mail: moschini@iastate.edu.

Gabriel E. Lade is Assistant Professor, Department of Economics, Macalester College, St. Paul, Minnesota 55105. E-mail: glade@macalester.edu.

For questions or comments about the contents of this paper, please contact GianCarlo Moschini, moschini@iastate.edu.

Geographical Indications and Welfare: Evidence from the US Wine Market †

Raj Chandra, GianCarlo Moschini, and Gabriel E. Lade *

Abstract. A systematic component of wine quality is believed to depend on the geo-climatic factors of its production conditions. This belief has long been a motivation for the development of geographical indications for wines. In the United States, American Viticulture Areas (AVAs) represent the most common geographic identifier firms use to differentiate their products. We estimate a discrete-choice model of US wine demand to study consumers' valuation of the geographic origin of wine, and the market and welfare impacts of AVAs and other US appellations for wine. Specifically, we develop a two-level nested-logit choice model featuring many wine products and characteristics—including wine type, brands, and varietals, in addition to geographic origin—and estimate it using Nielsen Consumer Panel data over the 2007–2019 period. We find significant welfare gains from information about the geographic origin of wines. Over the period of interest, the welfare gain attributable to US geographic origin designation is estimated at about $5.13 billion, with wine producers and retailers capturing approximately 77% of this surplus. Virtually all of consumer welfare gains are due to product differentiation and increased product variety enabled by information about the wine origin.

Keywords: American Viticulture Areas, geographic indications, nested logit, product differentiation, wine demand, willingness-to-pay, welfare.

JEL Codes: D82, O34, Q11, Q18

First Draft : 22 October 2021

Revision: 23 June 2023

* Raj Chandra, currently at Amazon, worked on this paper as part of her dissertation while a Ph.D. student at Iowa State University. GianCarlo Moschini is in the Center for Agricultural and Rural Development and the Department of Economics, Iowa State University. Gabriel E. Lade is in the Department of Economics, Macalester College.

† Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1. Introduction

Firms have long faced the thorny issue of supplying the product attributes sought by consumers in contexts inherently affected by asymmetric information. Market solutions for this problem, which rely on reputation-based mechanisms (Akerlof 1970; Klein and Leffler 1981; Shapiro 1983), involve the concept of brands—the use of commercial names, symbols, designs, or marks that make the specific product sold by a firm distinctive in the eyes of consumers (Farquhar 1989). Branding investments are typically firm-specific and proprietary, are legally rooted in the trademarks system, and the brand equity they make possible is among firms' most valued assets in a modern economy (Bronnenberg, Dubé, and Syverson 2022). A prominent and unique feature of the agricultural and food industry, however, is that firm-specific branding activities are often augmented by collective efforts pertaining to the certification of origin of the product by using so-called geographical indications (GIs).

GIs are rooted in France's pioneering development of "*appellations*" for wines (Stanziani 2004). Other European countries followed this approach, which eventually formed the blueprint for the European Union's general policy on the protection of denominations of origin (Meloni and Swinnen 2013). Recognition of GIs as a distinct form of intellectual property by the TRIPS Agreement of the World Trade Organization (Moschini 2004) further supported their spread beyond Europe. The GIs' logic of emphasizing a good's geographic origin is predicated on the belief that geography is inextricably linked to the quality of the good because of unique geo-climatic production conditions and local production processes, as captured by the notion of "terroir."[1] Credible certification of such an origin, therefore, can contribute substantially to alleviating asymmetric information problems (Moschini, Menapace, and Pick 2008; Menapace and Moschini 2012; Mèrel, Ortiz-Bobea, and Paroissien 2021).

GIs were first developed for, and are particularly prominent in, the wine market. Appellation-of-origin labels convey information about climate and soil conditions where grapes were grown, which are believed to be associated with important sensory qualities of wine (Ashenfelter 2008;

[1] Castriota (2020, pp. 62–64) provides some background on this oft-used (and occasionally misused) concept.

Ashenfelter and Storchman 2010). As with other food products, wine consumers have incomplete information about many of the intrinsic attributes of wine that they value, and must make decisions based on the limited observable information disclosed on products' labels. This information typically includes the brand, variety, vintage, and geographic origin of the wine. American viticulture areas (AVAs) are the primary identifiers US wine producers use to differentiate their products geographically. Administered by the federal government, the AVA program currently recognizes 267 areas across the United States. Other recognized appellations of origin that can be used on wine labels include the state (e.g., California) and county (e.g., Amador). Whereas the rapid diffusion of these certification of origin labels attests to the interest this collective certification has received, empirical evidence as to the extent of the actual impact of AVAs and other US designations of origin for wine is limited.

In this article, we study consumers' valuation of US appellations of origin in the US wine market. We take it as given that the US implementation of GIs provides a credible instrument for consumers to ascertain the origin of the wines they buy. We, therefore, ask how much the origin per se is valued over and above other desirable quality attributes wine may possess. We do so by estimating a structural model of wine demand using a large data set for at-home wine consumption from Nielsen Consumer Panel data. The specification of this model is rooted in the discrete-choice framework, whereby the products in the consumers' choice sets are explicitly defined by the attributes they possess. The basic premise—consistent with the findings in Combris, Lecoq, and Visser (1997)—is that the observable information concerning wine attributes, disclosed on wine labels, is credible and key to consumers' choices. Thus, the estimated demand model provides the vehicle to assess consumers' willingness to pay (WTP) for the product attributes explicitly included in the model. Given that, we ask how much the wine origin per se is valued by consumers over and above what is ascribed to the other desirable attributes wines may possess, and which we control for in the model.

The extent to which geography matters, for wines, has been the object of considerable research. In contrast with the approach taken in this paper, however, most of the existing work has utilized the hedonic price function methodology—see Outreville and Le Fur (2020) for a detailed review.

Rosen's (1974) seminal paper envisioned two empirical steps for the hedonic framework, the first to estimate the characteristics' implicit prices, and the second to assess the demand and supply determinants of these implicit prices. Subsequent work, however, has clarified the daunting identification challenges with this program (e.g., Brown and Rosen 1982; Ekeland, Heckman, and Nesheim 2002). As a result, as noted by Costanigro and McCluskey (2011), most applications are confined to the first stage of the hedonic price methodology. Empirical hedonic price functions, however, represent equilibrium relations, and cannot disentangle the separate impacts of demand and supply factors. The structural model developed in this article, on the other hand, offers the potential to uncover the consumers' valuation of wines' geographic origin.

The specific model that we implement is a two-level nested-logit demand model (Verboven 1996; Björnerstedt and Verboven 2016; Ciliberto, Moschini, and Perry 2019). The model is estimated using Nielsen Consumer Panel data from 2007 to 2019. The data cover the entire US market and include extensive and detailed product attributes. Given the intractably large number of distinct wines sold, we aggregate consumers' alternatives in their choice sets along some key dimensions while preserving the identity of the attributes of interest. The final dataset includes about 2,941 wine "products" encompassing 33 varietals produced in 79 domestic and foreign geographic regions.

The estimated model permits us to infer consumers' WTP for various wine characteristics. We find that consumers place a relatively high value on wines' geographic origins, distinct from the value consumers place on brand and varietal information. Estimated WTPs for US geographic origin turn out to be economically and statistically significant. We also document substantial heterogeneity across wines in consumers' WTP for AVAs, ranging from -$2.9 to $17.0 per bottle, relative to a generic California wine (the reference region). Appellations fetching the highest premiums are well-known AVAs, including Anderson Valley, Carneros, Chalk Hill, Knights Valley, Oakville, Santa Maria Valley, and Sonoma Valley.

Aggregating marginal WTP estimates provides a first-order approximation to measuring welfare gains attributable to US wine designations of origin. This method implies gains of about $1.95 billion from 2007 to 2019. The estimated model, however, permits a more thorough welfare assessment. By postulating an imperfectly competitive retailing supply side of the wine market, under the standard hypothesis of Bertrand-Nash price competition, we evaluate a counterfactual experiment where information about the geographic origin of US wine is stripped away, while all other product characteristics of the demand model are preserved, and equilibrium prices are adjusted. A comparison of the baseline and counterfactual models shows that, from 2007 to 2019, total consumer welfare was higher by $1.17 billion because of product differentiation enabled by US designations of origin. We also find that industry revenues are positively impacted by this product differentiation by $3.70 billion, for a total welfare gain attributable to information about the US origin of wine of $4.87 billion over the period of study.

Related literature include earlier work that probed the implicit rationale of GIs, that is, whether geo-climatic variables are related to quality metrics such as wine price (Gergaud and Ginsburgh 2008), ratings (Ashenfelter and Storchmann 2010), or vineyard sale prices (Cross, Plantinga, and Stavins 2011). As to whether geographic origin matters for consumer choice, even crude indicators such as country-of-origin labels appear to influence consumers' perceptions of a wine's quality (Veale and Quester 2008).[2] As noted, a distinctive feature of wine marketing is the coexistence of private brands with GIs, providing a role for both individual and collective reputation mechanisms. Landon and Smith (1997, 1998) find that they have a large price impact, a result confirmed in other settings (Hadj Ali and Nauges 2007; Castriota and Delmastro 2012; Frick and Simmons 2013; Marchini et al. 2014; Chandra and Moschini 2022). Schamel (2009) finds that both region and producer reputations have a significant and positive impact on wine prices (which has a larger price impact is context-specific). Costanigro, McCluskey, and Goemans (2010) find that winery or brand names contribute more to price premia than the appellation's reputation. Similarly, Cuellar and Claps (2013) find that the price premium for Napa over Sonoma Cabernet diminishes when they control for brand names. Keating (2020) shows that even sub-AVAs within Napa Valley have a substantial effect on wine prices.

[2] The importance of geographic origin has also been investigated for other food products, such as meat and dairy products (Awada and Yiannaka 2012; Loureiro and McCluskey 2000; Norris and Cranfield 2019; Slade, Michler, and Josephson 2019), and olive oil (Aprile, Caputo, and Nayga 2012; Menapace et al. 2011).

This article makes three main contributions. First, we add to the literature on wine economics by developing a structural framework, based on discrete-choice demand, suitable to evaluate the consumers' valuation of different wine characteristics. Second, we contribute to the empirical knowledge on determinants of wine demand by estimating the demand model using an extensive dataset on US wine purchases for home consumption. Beyond identifying the role of product attributes in wine demand, the model yields useful estimates of both product-level and aggregate wine demand elasticities. Third, we estimate the welfare effects of US wine appellations. These welfare measures contribute new evidence about GI's role in reducing asymmetric information and providing quality in food markets. Our WTP estimates, elasticities, and welfare estimates also provide important market insights for winemakers and other wine industry participants.

The rest of the paper is organized as follows. Section 2 provides a brief background on GIs in the United States wine industry. Section 3 discusses data and product definition. Section 4 develops models of wine demand and supply. Section 5 reports and discusses demand estimation results. Section 6 pairs our demand and supply models to simulate a counterfactual scenario and estimate welfare generated by the welfare gains attributable to US appellation of origin. Section 7 concludes.

2. Appellation of Origin and American Viticulture Areas

Most wine-producing countries utilize quality standards and regulations that affect the wine sector, including GI systems to identify and recognize distinct grape-growing regions, distinguishing regions based on their viticulture tradition, and labeling wines based on their origin (Meloni et al. 2019). Well-known early implementations of GI systems include the *Appellation d'Origine Contrôlée/Protégée* in France, and the *Denominazione di Origine Controllata/Garantita* in Italy. Whereas the European Union (and other countries, notably in South America) have opted to frame GIs within a *sui generis* system of intellectual property, the United States relies on the standard trademark system administered by the US Patent and Trademark Office (USPTO) (Menapace and Moschini 2012), whereby GIs are treated as certification marks or collective marks.

The Alcohol and Tobacco Tax and Trade Bureau (TTB), a unit of the US Department of the

Treasury, is the agency demanded to implement wine labeling regulations (TTB 2021). It oversees both mandatory and voluntary labeling programs for wines, including AVAs and other appellations of origin. An AVA is defined as a delimited grape-growing region with distinguishing features, a name, and a delineated boundary. The TTB designates and reviews all petitions to establish new or expand existing AVAs. Any producer can include an approved AVA or sub-AVA label on their wine so long as at least 85% of the volume of its grapes is grown and finished in the named viticulture area. Thus, US AVAs are purely a geographic origin certification, unlike GIs in the European Union, where production methods presumed to affect quality are also regulated (Menapace and Moschini 2012). Other common appellations of origin for wine, regulated by the TTB, include state and county designations.

Augusta, Missouri, was the first wine-growing region approved as an AVA by the TTB in June 1980, and the program has grown significantly since then. As of May 2023, there were 267 approved AVAs in the United States, 147 of which are in California. Approved AVAs include world-renowned wine-growing areas such as Napa Valley and Los Carneros (California), as well as lesser-known regions such as Shawnee Hills (Illinois) and Grand Valley (Colorado). Most AVAs are contained within a single state, though several multi-state designations exist. For instance, Washington and Oregon share the Columbia Gorge AVA, and the Lake Erie AVA extends across New York, Ohio, and Pennsylvania. AVAs need not have exclusivity over a territorial entity, and several AVAs overlap, especially in California.[3]

3. Data

Our primary data are Nielsen's Homescan Consumer Panel data for 2007 to 2019. The Consumer Panel data are a longitudinal panel of approximately 60,000 US households' grocery purchases.

[3] Some AVAs are large and can have heterogeneous geo-climatic conditions. Sub-AVAs are AVAs that are approved within the boundaries of an existing AVA. For example, the Napa Valley AVA spreads from southeast to northwest for approximately 30 miles and contains 16 sub-AVAs that are hotter in the north and east, and cooler in the south and west. More precisely defined appellations are believed to better convey the qualities specific to their area. The Upper Mississippi River Valley is the largest AVA, covering 29,914 square miles and stretching across four states (Minnesota, Wisconsin, Illinois, and Iowa).

The nationally representative panelists are geographically dispersed and demographically balanced. Nielsen stratifies their sample proportionately into 61 geographic areas, including 50 major markets and nine Census divisions. We restrict our analysis to the 50 'Scantrack' markets (these are similar to Metropolitan Statistical Areas used by the US Census Bureau).

Panelists provide a journal to Nielsen detailing purchases for every shopping trip they make, including the date, retailer/store, and total dollars spent on each trip. Panelists also provide detailed transaction information for each item, including each product's Universal Product Code (UPC), the quantity purchased, the per-unit price, and whether deals or coupons were available. Relevant to this study, Nielsen also provides shorthand descriptions for each product, from which we extract information on each wine's type, varietal, and geographic origin.[4]

3.1 Product Definition

The data include information for over 37,000 unique UPCs for wines purchased, with an average of 12,000 distinct UPCs purchased each year. Given the large number of products, aggregation is inevitable for demand estimation. To proceed, for the purpose of the model, we define a "product" as a unique combination of five characteristics: wine type, packaging size, brand name, varietal, and geographic origin, which we describe in further detail below.[5]

"Wine type," as defined here, comprises three categories: red wine, white wine, and specialty wine. The first two categories encompass all dry table and kosher table wine in the sample. The specialty wine group contains any wine other than table wine, as a specialty wine, mostly sparkling wine, but also blush and rosé. "Packaging size" distinguishes two categories: bottled wine, which pertains to wines sold in the standard 750 ml bottles, and bulk wine (any package containing more

[4] Appendix A reports market shares in quantity and sales value for many of the different product characteristics described in this section. Appendix C provides further details regarding how we extract wine characteristics from Nielsen's product descriptions and unique UPCs into products used in our demand estimation.

[5] Our data does not contain information in bottle vintage, a common challenge in the wine demand literature and key limitation if bottle vintage is correlated with other characteristics of interest.

than 750 ml). Next, we distinguish wines by their brand. Our data include 3,500 to 4,000 unique brands for any given year. Many brands have a very small market share, and none dominates the market. To simplify the analysis, we maintain brand distinctions for those with a market share of 0.1% or higher in the bottle category and 1% or higher in the bulk category. The top unique brands make up 69% of sales in the bottle category and 82% of sales in the bulk category. We jointly classify brands with smaller market shares as "other brands." Our procedure results in 154 distinct brands for bottled wine and 20 brands for bulk wine. Finally, we distinguish table wines by their "varietal." Similar to the case of brands, we observe many varietals. We keep unique varietal categories only for those with a market share greater than 0.2%. We aggregate other varietals with smaller market shares based on their grape type and import status. After aggregating smaller varietals, we have 14 red, 11 white, and eight specialty wine varietals.

Our final product attribute is "geographic origin." For domestic bottles, we extract 165 distinct geographic origins. Of these, 44 are state appellations, and the remaining 121 are AVAs or county appellations. We aggregate AVAs and county appellations with less than a 0.01% market share into an "other AVA," and we similarly aggregate state appellations with less than a 0.1% market share into "other states" category. After aggregation, we have 55 unique AVAs or county appellations and 12 state appellations. For bulk domestic wines, we identify 36 distinct geographic origins, of which 24 are AVAs or county, and 12 are state appellations.

The data also include purchases of foreign wines from 30 countries. Extracting accurate information about the specific foreign GI from information included in the UPC code turned out to be more difficult, however. Because the data only permit a very incomplete classification of foreign GIs, we elected to identify the geographic appellation of foreign wine simply in terms of the country of origin. We consider separately the top 10 countries, and the remaining 20 (which collectively account for less than 0.5% of total imports) are aggregated into "other countries." Whereas somewhat coarse, we believe that this classification of foreign wines provides an adequate control for the main focus of the analysis, which concerns the impact of GIs for US wines. Combining domestic and international designations, we have 79 bottled wine geographic origins and 48 bulk geographic origins.

Table 1 lists the category count for each of the product attributes. We classify every unique UPC in our data into one of 2,941 products over the 13 years of the sample. On average, we observe 306 unique products sold in a market (Scantrack market-year-quarter combination). **Figure 1(a)** displays market shares across three broad geographic-origin categories over time. In 2019 (the last year of the sample), 63% of sales were domestic wines with a state appellation (California is by far the largest), 27% were imported wines, and 10% belonged to an AVA or county appellation. The share of the latter increases slightly over time. **Figure 1(b)** shows that, while AVAs make up a small share of total sales, they fetch higher prices on average than other types of wines. In 2019, wines with AVA or country appellation labels had the highest average price ($11/bottle) compared to imported wines ($7/bottle), and state appellation wines ($4.20/bottle).

3.2 Projection Weights

Our data report transactions at the trip-UPC-household level. Nielsen assigns every panelist a sampling weight to project nationally representative purchases. We use the sampling weights to calculate the quantities and expenditures for each wine purchase in the data. We also use the weights to calculate product shares, prices, and market size for our demand estimation.[6] This allows us to more readily extrapolate our welfare calculations to national estimates in our counterfactual exercises.

3.3 Market Definition and Market Size

For the purpose of estimating the model, the unit of observation is a "market." Following customary procedures, a market is identified as a specific US region at a given period of time. Specifically, a market in our context is a Scantrack market-year-quarter combination. Having defined the scope of the analysis over the 50 Scantrack markets, as noted earlier, the 52 quarters over the period 2007 to 2019 provide a total of 2,600 unique markets. We note significant variation

[6] There are conflicting views on the use of weights in regression analysis. Solon, Haider, and Wooldridge (2015) suggest that weighted regressions are often unnecessary if we are concerned about causal effects of a treatment. In our setting, weights are necessary when presenting national summary statistics. For internal consistency, we also use weights in the demand model estimation. In any case, we verify that parameters are very similar when the demand models are estimated without weights.

in choice sets across markets. Of 2,941 products, the average number of products available in a market is 306. The smallest observed choice set in a market is 36 products, and the largest is 639.

The model's outside option is best interpreted as all other alcoholic beverages, such as beer and spirits, and its measurement implicitly relies on the assumed "potential market size." Standard practice in empirical industrial organization is to assume that market size is some multiple of observed average sales in each market (e.g., Björnerstedt and Verboven 2016; Miller and Weinberg 2017). Because we are studying demand for alcoholic beverages, we adjust this definition to account for changes in adult population over time. Specifically, we first define the maximum per capita wine consumption in each region over time. We then assume the per capita potential market is 50% larger than this maximum value. Multiplying this per capita potential market by the annual population above age 21 yields our measure of the potential market size.[7]

We follow standard practice and express market shares in physical units, specifically 750 ml bottle equivalent units (based on the most common authorized standard of fill for wine). The price p_{jm} of any product j in market m is calculated by dividing the dollar sales by the number of bottle equivalent units sold of product j. We deflate all prices to 2019 dollars. Our final data contain 794,985 total observations at the product-market level.

4. Empirical Framework

Our strategy is first to estimate the consumers' demand model. This provides direct evidence of the impact and importance of various attributes on consumer demand. Having postulated a market-equilibrium condition, we then consider counterfactual simulations that permit a fuller assessment of the welfare impacts of AVAs.

4.1 Demand

We rely on a discrete-choice framework, with preferences defined in the characteristics space, to rationalize consumer wine purchasing behavior. We specify a two-level, nested-logit demand

[7] Robustness results, reported in Appendix B, show that the arbitrary definition of potential market size has very little impact on estimated demand parameters.

model that allows consumers' unobserved tastes to be correlated across products (Verboven 1996; Björnerstedt and Verboven 2016). The model defines product groups, which allows for more realistic substitution patterns across products than a multinomial logit model.

Figure 2 illustrates the structure of the model. The upper level consists of an outside option and all wine products. We divide the inside option into three subgroups based on wine type: red, white, and specialty wine. This structure maintains an appealing condition: if a given consumer's utility from a particular bottle of red wine is low, she is more likely to consider another red wine rather than a white or specialty wine. The two-level nested structure also maintains another appealing condition: switching to the outside option (other alcoholic beverages) is less likely than switching between wine types.

Consider a market with L consumers denoted by $i = 1,....,L.$ Each consumer chooses one of $J+1$ differentiated products, where $j = 0,.....J.$ Good 0 is the outside option (the no-purchase alternative). Following Berry (1994), the conditional indirect utility that consumer i receives from good j in market m is written as:

(1) $$U_{ijm} = \delta_{jm} + v_{ijm}$$

where δ_{jm} denotes the mean utility of product j that is common across all consumers within market m, and v_{ijm} is a random (to the econometrician) term capturing consumers' idiosyncratic preferences. We follow common practice and, without further loss of generality, the mean utility of the outside good is normalized to zero (i.e., $\delta_{0m} = 0.$).

Product-specific mean utilities depend on the product's characteristics, including price. Specifically, we model mean utilities as follows:

(2) $$\delta_{jm} = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_{jm} + \xi_{l[m]} + \xi_{t[m]} + \xi_{\tau[m]} + \xi_{jm}$$

where $\mathbf{x}_j$ is a vector of observable product characteristics (packaging size, type, varietal, geographic origin, and brand); p_{jm} is the price of product j in market $m;$ $\xi_{l[m]}$, $\xi_{t[m]}$, and $\xi_{\tau[m]}$

are region, year, and quarter fixed effects, respectively; and, ξ_{jm} denotes unobserved product characteristics specific to a market that are observed by consumers and sellers but unknown to the econometrician. The notation $l[m]$ denotes the location of the (Nielsen Scantrack) market m, $t[m]$ denotes the year of market m, and $\tau[m]$ denotes the quarter associated with market m.

Assumptions about the distribution of the idiosyncratic term ν_{ijm} are needed to make the model operational. If this term were drawn from an i.i.d. Type 1 extreme-value distribution, the multinomial logit model would attain. But this model is known to entail restrictive implications for the cross-price elasticities (ultimately due to its independence of irrelevant alternative property). To overcome this limitation, as noted, we postulate a two-level nested-logit formulation. We partition the choice set in market m into two mutually exclusive groups denoted by $g \in \{0,1\}$, where $g = 0$ represents the outside good and $g = 1$ represents inside goods. We further partition products in the inside group into three subgroups denoted by $h \in \{1,2,3\}$ that represent the three different wine types. Given that, the idiosyncratic unobserved error term is written as (Verboven 1996):

(3) $$\nu_{ijm} = \nu_{igm} + (1-\sigma_2)\nu_{ihgm} + (1-\sigma_1)\varepsilon_{ijm}$$

where ε_{ijm} is an i.i.d. Type 1 extreme value draw, and ν_{igm} and ν_{ihgm} follow the unique distributions such that $(1-\sigma_2)\nu_{ihgm} + (1-\sigma_1)\varepsilon_{ijm}$, and $\nu_{igm} + (1-\sigma_2)\nu_{ihgm} + (1-\sigma_1)\varepsilon_{ijm}$ are also distributed as Type 1 extreme value.

The structure of equation (3) ensures that ν_{ijm} are not i.i.d. so long as the nesting parameters are not zero. The nesting parameters σ_1 and σ_2 allow for correlation between unobserved preference components. Specifically, σ_1 allows for taste correlations between products within the same subgroups, and σ_2 for correlation between products in the same group. To be consistent with the axioms of utility maximization, these nesting parameters should satisfy $0 \le \sigma_2 \le \sigma_1 < 1$. A large σ_1 indicates that product demand is highly correlated across products within the same subgroup (wine type). When σ_2 is also high, there is additional correlation across products of all three wine types. If $\sigma_2 = \sigma_1$, consumers' preferences are equally correlated among wine products, and the subgroup distinction is immaterial. If $\sigma_1 = \sigma_2 = 0$, the model reduces to a simple multinomial logit specification.

Each consumer in every market m chooses the product j among the available choices that yields the highest utility u_{ijm}. Conditional on choosing product j, a consumer will buy one unit of product j. In any given market, the set of wine products in subgroup h of group g is denoted J_{hgm}. Given the assumed two-level nested-logit structure, the choice probability that a consumer i chooses product $j = 1,\ldots,J$ is:

(4) $$s_{jm} = \frac{\exp\left(\delta_{jm} / (1-\sigma_1)\right)}{\exp\left(I_{hgm} / (1-\sigma_1)\right)} \frac{\exp\left(I_{hgm} / (1-\sigma_2)\right)}{\exp\left(I_{gm} / (1-\sigma_2)\right)} \frac{\exp(I_{gm})}{\exp(I_m)},$$

where I_{hgm}, I_{gm} and I_m are "inclusive values" defined as (Björnerstedt and Verboven 2016):

(5) $$I_{hgm} = (1-\sigma_1)\ln \sum_{k \in J_{hgm}} \exp(\delta_{km} / (1-\sigma_1))$$

(6) $$I_{gm} = (1-\sigma_2)\ln \sum_{h \in (1,2,3)} \exp\left(I_{hgm} / (1-\sigma_2)\right)$$

(7) $$I_m = \ln\left(1 + \exp(I_{gm})\right).$$

Following Berry (1994) and Verboven (1996), and recalling the definition of mean utility δ_{jm}, the estimating equation can then be expressed as:

(8) $$\ln(s_{jm} / s_{0m}) = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_{jm} + \sigma_1 \ln(s_{jm} / S_{hgm}) + \sigma_2 \ln(S_{hgm} / S_{gm}) + \xi_{l[m]} + \xi_{t[m]} + \xi_{\tau[m]} + \xi_{jm}$$

where, for product j that belongs to subgroup h of group g (and here we have only one group of inside goods), S_{hgm} is the aggregate share of all products in subgroup h, and S_{gm} is the aggregate share of all products in group g. Thus, s_{jm}/S_{hgm} is the conditional share of wine product j within subgroup h; and, S_{hgm}/S_{gm} is the conditional share of subgroup h in the group of all wines. The term s_{0m} is the market share of the outside option, which satisfies $s_{0m} = 1 - S_{gm}$.

4.2 Identification

Equation (8) can be estimated using linear regression, provided an instrumental variable approach is implemented to correct for price, subgroup share, and group share endogeneity. We assume that profit-maximizing firms set wine prices and consider all differentiating characteristics of their products and those of competing products. Because we do not observe all relevant characteristics, however, there is potential correlation between product prices and the structural error term ξ_{jm}.[8] In the presence of this price endogeneity, ordinary least squares (OLS) estimation of the price coefficient α will be biased towards zero, and subsequent WTP and welfare estimates will also be biased.

We follow the standard approach and assume that product characteristics other than price are uncorrelated with the error term, conditional on the model's fixed effects. We use different instrumental variable strategies to account for price endogeneity.[9] Given the longitudinal nature of our data, we follow Nevo (2001) and estimate our demand model in two steps. In the first step, we introduce product fixed effects ξ_j and estimate the following equation:

[8] We note at this juncture that prices in Nielsen data are determined in two ways. If the panelists have shopped at a store where Nielsen also receives point-of-sale (POS) data, then panelists are not supposed to enter the price information. For such cases, the price reported is the average weighted price of the item that week in that store. If the panelist shops at an outlet where Nielsen does not receive POS data, panelists enter the price paid. Nielsen does not distinguish the price entered by the panelists and those imputed by Nielsen, as they assume all price information in the Consumer Panel data is accurate.

[9] Other potential sources of endogeneity include the conditional subgroup and group shares in equation (8). These group shares are systematically related to overall market shares, which are endogenous due to unobserved (to the econometrician) product characteristics.

(9) $\ln(s_{jm}/s_{0m}) = \xi_j - \alpha p_{jm} + \sigma_1 \ln(s_{jm}/S_{hgm}) + \sigma_2 \ln(S_{hgm}/S_{gm}) + \xi_{l[m]} + \xi_{t[m]} + \xi_{\tau[m]} + \xi_{jm}$

Product fixed effects capture observed and unobserved product characteristics that are constant across markets.

We estimate equation (9) using four sets of instruments. Following Berry, Levinsohn, and Pakes (1995), we use the so-called BLP instruments that are common in the industrial organization literature. Specifically, we leverage variation in choice sets across markets and count the number of competitive products for several categories. The categories include group (all wine products), subgroup (wine types), size, varietal, geographic origin, brand, wine type and varietal, wine type and size, wine type and geographic origin, and wine type and brand.

Second, as in Miller and Weinberg (2017), we create a product-level cost instrument. Specifically, we proxy distribution costs for every product using travel distances between production and sales regions.[10] Next, we construct instruments based on the competitiveness of the local retail store environment. We assemble store count data for every market from the County Business Pattern of the US Census and calculate two measures of each market's retail density: (*a*) "population-adjusted retail density" (i.e., the number of establishments per 1,000 residents); and, (*b*) "spatial area adjusted retail density" (i.e., the number of stores per square mile). Lastly, we use state excise taxes, which vary considerably across states, as an additional instrument. Wine excise tax data are collected from the Tax Foundation and the Federation of Tax Administrators.

In the second step regression, we recover marginal utilities of observed product attributes that are fixed over time and across regions by regressing the estimated product fixed effects, $\hat{\xi}_j$ from equation (9), on observed attributes. We are particularly interested in varietal, geographic origin, and brand attributes. The second step regression can be expressed as:

(10) $\hat{\xi}_j = \mathbf{x}_j'\boldsymbol{\beta} + e_j$

where the vector $\mathbf{x}_j$ contains a set of indicator variables for the observable attributes of product j

[10] See Appendix C.3 for more details on our instrument construction.

—wine type, varietal, container size, geographic origin (AVA for domestic wines, and foreign geographic origin for imported wines), and brand—and e_j is an error term capturing the impact of unobserved characteristics.

4.3 Supply

To close the model, we adopt the Bertrand-Nash price competition model widely used in the literature on consumer products (Gandhi and Nevo 2021), including applications to the wine market (Villas-Boas, Bonnet, and Hilger 2020). Competition occurs between brands, and individual brands may be associated with multiple products. Dropping the market identifier, m, for notational clarity, suppose there are B brands in the market, each of which produces some subset J_b of $j = 1,.....,J$ wines. The profits of brand b can be written as:

$$(11) \qquad \Pi_b = \sum_{j \in J_b} (p_j - mc_j) M s_j(\mathbf{p})$$

where $s_j(\mathbf{p})$ is the market share of product j, which is a function of the vector of prices $\mathbf{p}$ of all products in any given market; M is the size of the market; and, mc_j is the (constant) marginal cost of product j.

We assume that firms compete in prices. The optimality conditions for the maximization of the profit in (11), which implicitly define the best response functions of the Bertrand-Nash game, are:

$$(12) \qquad s_j(\mathbf{p}) + \sum_{i \in J_B} (p_i - mc_i) \frac{\partial s_i(\mathbf{p})}{\partial p_j} = 0$$

Let $\mathbf{S}(\mathbf{p})$ be the $J \times J$ matrix of substitution terms (i.e., with elements $S_{jk} \equiv \partial s_j(\mathbf{p}) / \partial p_k$) and define a $J \times J$ ownership matrix $\mathbf{H}$ with elements $\mathbf{H}_{ij} = 1$ if products i and j are sold by the same brand, and $\mathbf{H}_{ij} = 0$ otherwise.[11] Following Nevo (2001), the information from the substitution matrix and

[11] When defining products, we aggregate brands that do not meet the cutoff (top 69% for bottle and top 80% for bulk) as "other" brand. Because this group does not reflect ownership, however,

the ownership matrix can be combined in a $J \times J$ matrix $\mathbf{\Omega}(\mathbf{p})$, with elements $\Omega_{jk} \equiv -S_{jk} \times H_{jk}$. The Bertrand-Nash equilibrium conditions in equation (12) can then be expressed as:

$$(13) \quad s(\mathbf{p}) - \mathbf{\Omega}(\mathbf{p})\left[\mathbf{p} - \mathbf{mc}\right] = 0,$$

where $s(\mathbf{p})$ and $\mathbf{mc}$ are the vectors of market shares and marginal costs, respectively. Equation (13) implies the equilibrium markup relation $\mathbf{p} - \mathbf{mc} = \mathbf{\Omega}^{-1}s(\mathbf{p})$, which embeds the insight that markups depend on price elasticities and also on the product ownership structure (because firms internalize the effects of their price choices on the demand for other products they sell). Having estimated the demand model, in the tradition of the new empirical industrial organization, we exploit this equilibrium relation to back out the vector of (otherwise unobserved) marginal costs, which are needed in the counterfactuals to assess the welfare impacts.

5. Results

In this section, we present estimates from alternative demand models. We then derive and discuss relevant elasticities of demand, comparing them to previous work.

5.1 Demand Model Results

Table 2 presents estimation results for four specifications of our demand model. Columns 1 and 2 contain results for the two-level nested-logit specification model, the only difference being in the instrumental variables used in estimation. Specifically, column 1 uses all of the instruments discussed earlier (BLP instruments, product-cost variable, market retail density metrics, and excise taxes), whereas column 2 only uses BLP instruments. In addition to the results for the nested-logit model, for comparison purposes we also report the results for the simple multinomial logit specification, where $\sigma_1 = \sigma_2 = 0$. In column 4, the simple multinomial logit is estimated by OLS, whereas column 3 reports the results of the simple multinomial logit model estimated using all of the instrumental variables. All specifications also include quarter, year, region, and product fixed

when modeling brand's competition through the ownership matrix, we treat all brands in the group as independent.

effects.

Comparison of columns 3 and 4 for the simple logit model shows that the price coefficient is substantially smaller (in absolute value) when instrumental variables are not used, a finding consistent with the presence of price endogeneity (which, when not accounted for, biases estimates towards zero). A comparison of columns 1 and 2, for the nested logit model, shows that the use of the other instrumental variables we have constructed, beyond the BLP instruments common in empirical industrial organization, significantly impacts the estimated price coefficient in the expected direction. In particular, the price coefficient in column 1, our preferred model, is statistically significant and negative, and the nesting parameters have the expected sign and relative magnitude (recall that consistency with theory requires $0 \le \sigma_2 \le \sigma_1 < 1$). The estimated nesting parameters indicate a strong correlation within nests, suggesting that consumers consider products of the same wine type as closer substitutes. These results provide some support for the rationalization of the usefulness of the two-level nested-logit model, and indicate that the simple multinomial logit model is clearly inadequate.[12]

5.2 Product Elasticities

Following Björnerstedt and Verboven (2016), the price elasticity for any product j with respect to a change in the price of product i , in our two-level nested-logit model, is:

$$(14) \qquad \varepsilon_{ji} \equiv \frac{\partial s_j}{\partial p_i}\frac{p_i}{s_j} = -\alpha\left(\frac{1}{1-\sigma_1}D_{ji} - \left(\frac{1}{1-\sigma_1} - \frac{1}{1-\sigma_2}\right)\frac{s_i}{S_{hg}}D_{ji}^{SG} - \frac{\sigma_2}{1-\sigma_2}\frac{s_i}{S_g}D_{ji}^{G} - s_i\right)p_i$$

where D_{ji}, D_{ji}^{SG}, and D_{ji}^{G} are dummy variables with $D_{ji} = 1$ if $j = i$ ($D_{ji} = 0$ otherwise), $D_{ji}^{SG} = 1$ if j and i belong to the same subgroup ($D_{ji}^{SG} = 0$ otherwise), and $D_{ji}^{G} = 1$ if j and i belong to the same group ($D_{ji}^{G} = 0$ otherwise). Here, s_i / S_{hg} is the conditional share of product i in subgroup

[12] Appendix B explores the sensitivity of our baseline demand estimates to using alternative market size definitions, and using year trends rather than fixed effects. While alternative market size does not affect our results, using a more restrictive trend to control for linear changes in demand over time results in higher price sensitivity estimates. We also present first-stage estimates for all instrumental variables in Appendix B. Coefficients generally follow their expected signs.

hg, and s_i / S_g is the conditional share of product i in group g. Equation (14) reduces to an own-price elasticity when $j = i$.

The second panel of **Table 2** presents averages of own- and cross-price elasticities for all four models. For brevity, we focus on elasticity estimates from our preferred model in column 1, except to note that the simple logit model in column 4 produces inelastic own-price elasticity estimates on average, which would be inconsistent with profit-maximizing firms optimally setting prices. The mean own-price elasticity for our preferred nested logit model is -4.75. Thus, demand for individual wine products is highly elastic on average, as one would expect given the highly refined nature of the product space. We present two cross-price elasticity statistics: the average cross-price elasticity within a wine type, and the average cross-price elasticity across wine types. As expected, cross-price elasticities are larger (0.006) when we consider only products within the same wine type versus across different wine types (0.002). All average cross-price elasticities are positive, reflecting the substitutability of wine products. The average cross-price elasticity with respect to the outside good (other alcoholic beverages) is even smaller (0.0008).

5.3 Aggregate Demand Elasticity

The foregoing estimated demand elasticities pertain to the specific product definition used in this article and have a well-known interpretation within the maintained imperfectly competitive model of product differentiation. For some policy questions, however, students of wine demand may desire aggregate elasticity of wine demand, and indeed that is the focus of traditional—product-based as opposed to characteristic-based—models (Gallet 2007). To derive such an aggregate elasticity from our disaggregated demand model, let $s(\mathbf{p})$ denote the share of all wines (inside goods) in the potential market. Consider scaling the prices of all wine products by $k > 0$ (i.e., $p_\ell \to kp_\ell, \forall \ell \neq 0$— the price of the outside good is held constant). The aggregate wine demand elasticity is then defined as:

$$(15) \qquad \varepsilon \equiv \frac{\partial s(k\mathbf{p})}{\partial k} \frac{k}{s(k\mathbf{p})}\bigg|_{k=1} \quad \rightarrow \quad \varepsilon = \sum_j \sum_i \frac{\partial s_j}{\partial p_i} \frac{p_i}{s}.$$

where $s(k\mathbf{p}) \equiv \sum_{j=1}^{n} s_j(k\mathbf{p})$ is the aggregate share of all wine in the potential market. Hence, the aggregate elasticity of interest can be computed from the individual-product elasticities ε_{ji} reported in equation (14), as follows:[13]

$$(16) \qquad \varepsilon = \frac{1}{s}\sum_{j=1}^{n}\sum_{i=1}^{n} s_j \varepsilon_{ji} \ .$$

The estimated average aggregate wine demand elasticity, as per the foregoing definition, is also reported for each model in **Table 2**. For our preferred model in the first column, this elasticity value is –0.53. This indicates that, in the aggregate and on average, wine demand is inelastic. This value appears reasonable, and comparable with elasticity estimated obtained by rather different models.[14]

6. Willingness-to-Pay for Wine Attributes and US Wine GIs

We now turn to the second step of our demand model, estimating consumers' WTP for fixed wine characteristics, including AVAs. Beyond their direct interest, these estimates, together with a simulated counterfactual equilibrium, permit an assessment of the welfare impact of AVAs and other US appellations of origin. Throughout, we use estimates from our preferred demand model as presented in column 1 of **Table 2**.

6.1 Consumers' Willingness-to-Pay for Geographic Origin

Table 3 presents results from the second step regression. The first and second columns present coefficients and standard errors of the estimated marginal utilities that are the direct output of the regression in equation (10). The third and fourth columns report WTP estimates and their standard

[13] Note that this formulation is independent of how shares are modeled—our nested-logit specification simply determines the parametric structure of the individual (product-level) elasticities. For the simpler multinomial logit model, the aggregate elasticity defined here reduces to the analytic expression derived by Davis and Wilson (2005).

[14] For example, Hart and Alston (2020), using a Rotterdam model, estimate US aggregate wine demand for the period 1997–2016 to be –0.45.

errors. We calculate WTP as the ratio of the marginal utility for each characteristic to the marginal utility of price, estimated in the first stage of the demand model (e.g., Train 2009, p. 39). Two features of the estimates are worth noting. First, all WTP estimates are in real terms (2019 dollars). Second, all characteristics are product categories represented in our model by dummy variables. Choosing a "base" (excluded characteristic) for every category is therefore necessary. For example, every product is either a specialty, white, or red wine. For this case, we exclude the red wine indicator variable in our model and interpret all WTP estimates relative to a generic red wine. Each panel in **Table 3** presents results for a different exhaustive category, indicating the reference characteristic in the panel description.

Panel A shows that consumers are willing to pay, ceteris paribus, around $2.40 more per bottle, for bottled wine relative to bulk wine. Consistent with previous research, we find in panel B that consumers value red wine more than white wine or specialty wines (Schamel 2006; Schamel and Anderson 2003). Panels C, D, and E show results for several varietals separately for red, white, and specialty wines. Again, it is important to note that these WTPs are to be interpreted relative to the arbitrarily chosen reference varietal. It is apparent that, for red wine, consumers value Cabernet Sauvignon and Pinot Noir significantly more than Merlot (the reference varietal). Zinfandel and Petit Syrah are also valued more than Merlot, but these estimates are not statistically significant. Among white wines, Chardonnay, Moscato, Pinot Grigio, and Viognier all appear preferred to the Sauvignon Blanc (the reference varietal). Among specialty wines, the highest WTP relative to the reference attribute (blush rosé) is for sparkling wines.

Panel F reports the estimated marginal utilities and WTPs for our main object of interest, US geographical origin. The reference category is a bottle of wine with a generic California label. The AVAs with the largest estimated marginal value, relative to the California reference label, are from well-known wine-producing regions that fetch high price premiums: Anderson Valley ($9.83/bottle), Carneros ($9.27/bottle), Chalk Hill ($15.4/bottle), Knights Valley ($12.3/bottle), Oakville ($17/bottle), Santa Maria Valley ($9.03/bottle), and Sonoma Valley ($12.0/bottle). Other well-known AVAs (Alexander Valley, Dry Creek Valley, Livermore Valley, Napa Valley, Russian River Valley, and Rutherford) have estimated WTPs in the $6.00–7.00/bottle range. Few AVAs have negative WTPs (relative to the reference California).

The ranking of our estimated WTPs for AVA characteristics broadly aligns with those presented in earlier analyses (Bombrun and Sumner 2003; Schamel 2006; Costanigro, McCluskey, and Mittelhammer 2007; Gustafson, Lybbert, and Sumner 2016). Note, however, that our estimates are not directly comparable because the aforementioned prior works rely on hedonic price analysis, whereas we have estimated a structural demand model. As discussed earlier, hedonic price equations represent equilibrium relationships that are typically unsuited to recover purely demand-side elements such as consumers' WTPs. Furthermore, previous work often focuses on larger growing areas than individual AVAs. For example, three California regions included in Schamel (2006) are the South Coast, Napa Valley, and the Central Coast. These three areas include 10 distinct AVAs in our sample.

Panel G shows WTP estimates for state designations explicitly included in the model. Relative to the reference state California, most states' estimated WTPs are not statistically different from zero, with the exception of New York, which displays a significantly lower WTP. Panel H reports the estimated marginal utilities and WTPs for the foreign origin of wines (again, the reference is US wine with the generic California label). France ($3.49/bottle) and Italy ($1.59/bottle) are the foreign countries with the largest premium, whereas South Africa, Australia, and Chile appear characterized by lower WTPs relative to the reference (California).

The estimated marginal WTPs enable a first assessment of the welfare impact of US GIs. Multiplying the WTP estimates in **Table 3** by the observed number of bottles for each label, and summing the values for all US appellations, results in an estimated aggregate surplus of about $1.95 billion. This simple approach, however, is too coarse and does not consider how consumers and producers would react in a world where AVAs and other US appellations did not exist. Essentially, it assumes that sales of all wines would be unaffected by the removal of the AVA system, and the price of wines without AVA information would remain the same. Whereas the $1.95 billion figure may provide a useful benchmark, the structural nature of the model we estimate permits a more nuanced welfare assessment.

6.3 Counterfactual Scenarios and Consumer Welfare

The estimated demand model characterizes consumer behavior given the availability of information about the geographic origin of US wines—as well as other relevant attributes, such as prices, wine type, varietals, and brands. This estimated baseline can be compared to a counterfactual situation where all other attributes are unchanged, but information about the geographic origin of US wines is stripped away. Because modifying products' attributes affects product demand, in this context it is essential to also estimate the equilibrium prices and market shares that would prevail in the counterfactual situation. Having estimated the demand model, and explicitly articulated the supply side, Nash equilibrium conditions can be used to compute equilibrium prices and shares in the counterfactual, thereby permitting a calculation of internally consistent welfare measures.

The thought experiment addressed by the counterfactual simulations of this section concerns a world where information about the geographic origin of US wines does not exist. The simplest way to implement this experiment would be by removing all products with explicit AVA designations from the consumers' choice sets. Such a naïve procedure, however, is obviously uninteresting: even without a system of certification marks for the origin of wines, wineries would still be able to market their wines by using their own brands, a feature captured by our model. Thus, rather than reducing the size of consumers' choice sets, in our counterfactual scenario we modify the characteristics of all products by removing the information about their geographic origin, while keeping all other attributes that define a "product" in our setting. Thus, both the baseline and the counterfactual scenarios have the same number of alternatives in consumers' choice sets.[15]

The counterfactual experiments amount to solving the equilibrium model under alternative assumptions about the (exogenous) characteristics of the products. Recall that the estimated mean

[15] This characteristics modification may create "duplicate" products in our data. For example, if a specific brand markets the same varietal (e.g., pinot noir) from two different regions (e.g., Russian River Valley and Mendocino), then after stripping the information about the geographic origin these two products would look the same. We assume that duplicates, if they arise, are all viable and are part of the counterfactual choice set.

utility in the baseline model (i.e., with current AVA characteristics) in the given market is:

(17) $$\hat{\delta}_{jm} = \hat{\xi}_j - \hat{\alpha} p_{jm} + \hat{\xi}_{l[m]} + \hat{\xi}_{t[m]} + \hat{\xi}_{\tau[m]} + \hat{\xi}_{jm}$$

Product attributes that are fixed over time and across regions enter through the estimated product fixed effects $\hat{\xi}_j$. The impact of these variables is recovered from the second-step regression $\hat{\xi}_j = \mathbf{x}_j'\hat{\boldsymbol{\beta}} + \hat{e}_j$, where the vector $\mathbf{x}_j$ denotes the set of indicator variables for the attributes of product j—wine type, varietal, container size, geographic origin, and brand.

Changing the information about the geographic origin of wines directly affects the vector of attributes $\mathbf{x}_j$, which can be readily reflected in the counterfactual product fixed effects. Furthermore, in our counterfactual we also want to reflect the Akerlof (1970) insight that, in the pooling equilibrium where qualities cannot be discerned, consumers' expectation of the average quality may change. This is arguably important for questions pertaining to the geographic origin of wines. For example, even though the consumer may not be able to tell where the California wine comes from exactly, knowing that wines from Napa and Sonoma (say) are part of the residual California label of the counterfactual is likely to affect beliefs about the average quality of the wine. In other words, if $\tilde{\mathbf{x}}_j$ denotes the vector of attributes of the counterfactual, product fixed effects in the counterfactual are constructed as $\tilde{\xi}_j = \tilde{\mathbf{x}}_j'\overline{\boldsymbol{\beta}} + \hat{e}_j$. Here, $\overline{\beta}_k = \hat{\beta}_k$ for all the attributes other than US appellations (wine type, varietals, brands, and foreign country of origin). For variables coding for the US geographic origin of wines, however, $\overline{\beta}_k$ is the weighted average of the estimated marginal utility parameters associated with this geographic origin.[16]

Having characterized the counterfactual product fixed effects as in the foregoing paragraph, the mean utility terms in the counterfactual are:

(18) $$\tilde{\delta}_{jm} = \tilde{\xi}_j - \hat{\alpha} p_{jm} + \hat{\xi}_{l[m]} + \hat{\xi}_{t[m]} + \hat{\xi}_{\tau[m]} + \hat{\xi}_{jm}$$

[16] The weights, for the purpose of this average, are conditional national product market shares (that is, products market shares within domestic wines).

With these counterfactual mean utilities, the estimated structural demand model can predict the corresponding market shares $\tilde{s}_j(\mathbf{p})$. Such counterfactual shares depend on the vector of prices, which will itself be affected in the given scenario. Thus, we compute the vector of prices $\tilde{\mathbf{p}}$ that clears the market, simulating the corresponding Bertrand-Nash equilibrium conditions. Specifically, the counterfactual price vector satisfies:

$$(19) \qquad \tilde{\mathbf{p}} = \mathbf{mc} + \mathbf{\Omega}(\tilde{\mathbf{p}})^{-1}\tilde{\mathbf{s}}(\tilde{\mathbf{p}})$$

We use a fixed-point algorithm to solve the system of equations in equation (19). The vector of marginal costs, as estimated out from the equilibrium conditions in the baseline model, is held fixed, and the simulation iterates until the left- and right-hand sides of equation (19) converge.

The predicted mean utilities, evaluated at the appropriate price vector (the observed $\mathbf{p}$ for the baseline and the simulated $\tilde{\mathbf{p}}$ for each counterfactual), are then used to calculate the inclusive value for the baseline ($\hat{I}_m$) and counterfactual scenario ($\tilde{I}_m$). The change in consumer surplus between the baseline and a given counterfactual scenario for each market m, can be expressed as $V_m \equiv (\hat{I}_m - \tilde{I}_m)/\hat{\alpha}$. Multiplying these V_m terms by the potential market size M_m for the corresponding market m and summing over all markets yields an estimate of the total consumer surplus attributable to the availability of AVA attributes in wine products.

In the foregoing calculation, there are two channels for the information about the geographic origin of US wines to affect consumer surplus. First, and more directly, consumers can enjoy a larger variety of wines in the baseline, relative to the counterfactual, as the market pivots from a pooling equilibrium to a separating equilibrium. Second, the two settings (baseline and counterfactual) entail different equilibrium prices, which also affect consumer surplus. **Table 4** summarizes the results of the counterfactual experiment. When all US wines are stripped of their geographic origin information, the average price for US wine in the counterfactual scenario is \$4.90/bottle. This represents an increase of about 13% of the prices of California wines (the reference base), but a large decrease in the price of US premium wines (the prices of AVA wines declines by 57%). What most impacts consumer welfare, however, is the increased variety and product differentiation brought about by the information on US wines' geographic origin. Consumer surplus in the baseline, over the sample period, is \$1.18 billion larger than in the counterfactual, and most of this consumer gain is due to what we have labeled the "variety effect."

The ability of sellers to differentiate their wine products by geographic origin impacts their revenue as well. The estimated equilibrium prices for both baseline and counterfactuals allow us to evaluate the magnitude of this effect. Specifically, observed prices and counterfactual prices, along with the predicted mean utilities from equations (17) and (18), can be used to compute product market shares under the baseline ($\hat{s}_{jm}$) and counterfactual scenarios ($\tilde{s}_{jm}$). The change in industry revenue due to increase product differentiation enabled by information on the geographic origin of US wines, in any given market, is then estimated as $\Delta R_m \equiv M_m \cdot \left(\sum_j \hat{s}_{jm} p_{jm} - \sum_j \tilde{s}_{jm} \tilde{p}_{jm} \right)$. The total revenue change for the entire industry over the sample period is the summation of these terms over all markets, that is $\sum_m \Delta R_m$. The results of these calculations, also reported in **Table 4**, show that the ability to differentiate US wines based on their geographic origin—AVAs, state, and other appellations—leads to a large increase in wine revenue, \$3.95 billion over the study period.

The total welfare gain of \$5.13 billion is significant in economics terms. Over the period of analysis, this corresponds to about 12.9% of the estimated houseold expenditures on wine for at-home consumption.

7. Conclusion

The specific attributes of individual wines has long been linked, inter alia, to the soil and climatic conditions of the producing region. To credibly communicate these characteristics to consumers, over and above the firms' own brand labels, a system of GIs has emerged, first in Europe and then elsewhere around the globe. In the United States, AVAs represent the most prominent implementation of the GI concept.

The widespread diffusion of AVAs since their introduction in 1980 attests to the marketing attractiveness, to the wine industry, of using geographic origin information. Their number has steadily increased over the last few decades, even though establishing an AVA can be a lengthy and costly process (Walker 2006). That growers and wineries invest time and money to establish AVAs suggests that they believe in long-term gains to promote consumer awareness of their wines' geographic origin. Indeed, anecdotal evidence from industry observers suggests the possibility of large wine price premiums associated with AVAs (Chalk 2020). This study provides empirical backing for this perspective. Results from our structural wine demand model are broadly supportive of the role of GIs as an important instrument for product differentiation in the US wine market. We find that AVAs and other US appellations have an economically and statistically meaningful impact on wine demand and are associated with large welfare gains. In particular, our results indicate that the availability of information about the geographic origin of US wines leads to a sizeable increase in consumer surplus, and even larger increases in industry revenue.

Conventional wisdom holds that GIs have represented the hallmark of "old world" marketing strategies for wines, while the "new world" has privileged other avenues, including greater interest in technological innovation and a focus on varietals (Castriota 2020). Whereas the international landscape of the wine market is complex, with clear cross-country differences in consumer attitudes, production practices, and government regulations, our results suggest that the merits of GIs are not confined to their ancestral European home. It seems that US consumers recognize a clear value in AVA and other US appellations of origin. This is not surprising. As much as it may be challenging for consumers to deal with a large number of GIs on wine labels, it is a fact that these represent a more concise message relative to the panoply of branded individual products on the market. Insofar as key specific geo-climatic factors of the production locality matter for wine characteristics of interest to consumers, GIs can provide an effective informative signal that, consistent with Mèrel, Ortiz-Bobea, and Paroissien's (2021) analysis of the French experience, may go a long way to alleviating the informational problems that are pervasive in the highly differentiated wine industry.

References

Akerlof, G. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." ***The Quarterly Journal of Economics*** 84(3):488–500.

Aprile, M.C., V. Caputo, and R.M. Nayga Jr. 2012. "Consumers' Valuation of Food Quality Labels: The Case of the European Geographic Indication and Organic Farming Labels." ***International Journal of Consumer Studies*** 36(2):158–165.

Ashenfelter, O. 2008. "Predicting the Quality and Prices of Bordeaux Wine." ***The Economic Journal*** 118(529):F174–F184.

Ashenfelter, O., and K. Storchmann. 2010. "Using Hedonic Models of Solar Radiation and Weather to Assess the Economic Effect of Climate Change: The Case of Mosel Valley Vineyards." ***The Review of Economics and Statistics*** 92(2):333–349.

Awada, L., and A. Yiannaka. 2012. "Consumer Perceptions and the Effects of Country of Origin Labeling on Purchasing Decisions and Welfare." ***Food Policy*** 37(1):21–30.

Berry, S.T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." ***The RAND Journal of Economics*** 25(2):242–262.

Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." ***Econometrica*** 63(4):841–890.

Björnerstedt, J., and F. Verboven. 2016. "Does Merger Simulation Work? Evidence from the Swedish Analgesics Market." ***American Economic Journal: Applied Economics*** 8(3):125–64.

Bombrun, H., and D.A. Sumner. (2009). "What Determines the Price of Wine?" ***AIC Issues Brief***, 18, pp. 1-6.

Bronnenberg, B.J., J.P. Dubé, and C. Syverson. 2022. "Marketing Investment and Intangible Brand Capital." ***Journal of Economic Perspectives*** 36(3):53–74.

Brown, J.N., and H.S. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." ***Econometrica*** 50(3):765–768.

Castriota, S. 2020. ***Wine Economics***. Cambridge: MIT Press.

Castriota, S. and Delmastro, M., 2012. Seller reputation: Individual, collective, and institutional factors. ***Journal of Wine Economics***, 7(1), pp. 49-69.

Chalk, A. 2020. "How Much is an AVA Worth? How about a 25% Price Increase on Every Bottle of Wine." ***The Chalk Report,*** March 6, 2020.

Chandra, R. and Moschini, G., 2022. Product differentiation and the relative importance of wine attributes: US retail prices. ***Journal of Wine Economics***, 17(3), pp.177-208.

Ciliberto, F., G. Moschini, and E.D. Perry. 2019. "Valuing Product Innovation: Genetically Engineered Varieties in US Corn and Soybeans." ***The RAND Journal of Economics*** 50(3):615–644.

Combris, P., S. Lecocq, and M. Visser. 1997. "Estimation of a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter?" ***The Economic Journal*** 107(441):390–402.

Correia, S. 2016. "REGHDFE: Estimating Linear Models with Multi-way Fixed Effects." ***2016 Stata Conference*** (No. 24). Stata Users Group.

Correia, S. 2018. "IVREGHDFE: Stata Module for Extended Instrumental Variable Regressions with Multiple Levels of Fixed Effects." ***Boston College, Department of Economics***, Boston, MA.

Costanigro, M., J.J. McCluskey, and R.C. Mittelhammer. 2007. "Segmenting the Wine Market Based on Price: Hedonic Regression when Different Prices Mean Different Products." ***Journal of Agricultural Economics*** 58(3):454–466.

Costanigro, M., J.J. McCluskey, and C. Goemans. 2010. "The Economics of Nested Names: Name Specificity, Reputations, and Price Premia." ***American Journal of Agricultural Economics*** 92(5):1339–1350.

Costanigro, M., and J.J. McCluskey. 2011. "Hedonic Price Analysis in Food Markets." In: ***The Oxford Handbook of the Economics of Food Consumption and Policy***, 152–180. Oxford: Oxford University Press.

Cross, R., A.J. Plantinga, and R.N. Stavins. 2011. "What is the Value of Terroir?" ***American Economic Review*** 101(3):152–56.

Cuellar, S.S., and M. Claps. 2013. "Differential Effects of Brand, Ratings and Region on Willingness to Pay: A Hedonic Price Approach." ***Journal of Wine Research*** 24(2):138–155.

Davis, D.D., and B.J. Wilson. 2005. "Differentiated Product Competition and the Antitrust Logit Model: An Experimental Analysis." ***Journal of Economic Behavior & Organization*** 57(1):89–113.

Ekeland, I., J.J. Heckman, and L. Nesheim. 2002. "Identifying Hedonic Models." ***American Economic Review*** 92(2):304–309.

Farquhar, P.H., 1989. "Managing Brand Equity." ***Marketing Research*** 1(3):24–33.

Frick, B. and Simmons, R., 2013. The impact of individual and collective reputation on wine prices: empirical evidence from the Mosel valley. ***Journal of Business Economics***, 83, pp.101-119.

Gallet, C.A., 2007. “The Demand for Alcohol: A Meta-analysis of Elasticities.” ***Australian Journal of Agricultural and Resource Economics*** 51(2):121–135.

Gandhi, A., and A. Nevo. 2021. “Empirical Models of Demand and Supply in Differentiated Products Industries.” In ***Handbook of Industrial Organization*** Vol. 4, 63–139. Elsevier.

Gergaud, O., and V. Ginsburgh. 2008. “Natural Endowments, Production Technologies and the Quality of Wines in Bordeaux. Does Terroir Matter?” ***The Economic Journal*** 118(529):F142–F157.

Gustafson, C.R., T.J. Lybbert, and D.A. Sumner. 2016. “Consumer Sorting and Hedonic Valuation of Wine Attributes: Exploiting Data from a Field Experiment.” ***Agricultural Economics*** 47(1):91–103.

Hart, J., and J.M. Alston. 2020. “Evolving Consumption Patterns in the US Alcohol Market: Disaggregated Spatial Analysis.” ***Journal of Wine Economics*** 15(1):5–41.

Hadj Ali, H., and C. Nauges. 2007. “The Pricing of Experience Goods: The Example of En Primeur Wine.” ***American Journal of Agricultural Economics*** 89 (1): 91–103.

Klein, B., and K.B. Leffler. 1981. “The Role of Market Forces in Assuring Contractual Performance.” ***Journal of Political Economy*** 89(4):615–641.

Keating, G.B. 2020. “An Empirical Analysis of the Effect of Sub-divisions of American Viticultural Areas on Wine Prices: A Hedonic Study of Napa Valley.” ***Journal of Wine Economics*** 15(3):312–329.

Landon, S., and C.E. Smith. 1997. “The Use of Quality and Reputation Indicators by Consumers: The Case of Bordeaux Wine.” ***Journal of Consumer Policy*** 20(3):289–323.

———. 1998. “Quality Expectations, Reputation, and Price.” ***Southern Economic Journal*** 64(3):628–647.

Loureiro, M.L., and J.J. McCluskey. 2000. “Assessing Consumer Response to Protected Geographical Identification Labeling.” ***Agribusiness: An International Journal*** 16(3):309–320.

Meloni, G., and J. Swinnen. 2013. “The Political Economy of European Wine Regulations.” ***Journal of Wine Economics*** 8(3):244–284.

Meloni, G., K. Anderson, K. Deconinck, and J. Swinnen. 2019. “Wine Regulations.” ***Applied Economic Perspectives and Policy*** 41(4):620–649.

Menapace, L., G. Colson, C. Grebitus, and M. Facendola. 2011. “Consumers’ Preferences for Geographical Origin Labels: Evidence from the Canadian Olive Oil Market.” ***European Review of Agricultural Economics*** 38(2):193–212.

Menapace, L., and G. Moschini. 2012. “Quality Certification by Geographical Indications, Trademarks and Firm Reputation.” ***European Review of Agricultural Economics*** 39(4):539–566.

Mérel, P., A. Ortiz-Bobea, and E. Paroissien. 2021. “How Big is the ‘Lemons’ Problem? Historical Evidence from French Wines.” ***European Economic Review*** 138.

Miller, N.H., and M.C. Weinberg. 2017. “Understanding the Price Effects of the MillerCoors Joint Venture.” ***Econometrica*** 85(6):1763–1791.

Marchini, A., Riganelli, C., Diotallevi, F. and Paffarini, C., 2014. Factors of collective reputation of the Italian PDO wines: An analysis on central Italy. ***Wine Economics and Policy***, 3(2), pp. 127-137.

Moschini, G. 2004. “Intellectual Property Rights and the World Trade Organization: Retrospect and Prospects.” In G. Anania, M. Bohman, C. Carter, and A. McCalla (Eds.) ***Agricultural Policy Reform and the WTO: Where Are We Heading?*** 474–511. Northampton: Edward Elgar Publishing.

Moschini, G., L. Menapace, and D. Pick. 2008. “Geographical Indications and the Competitive Provision of Quality in Agricultural Markets.” ***American Journal of Agricultural Economics*** 90(3):794–812.

Nevo, A. 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” ***Econometrica*** 69(2):307–342.

Norris, A., and J. Cranfield. 2019. “Consumer Preferences for Country-of-Origin Labeling in Protected Markets: Evidence from the Canadian Dairy Market.” ***Applied Economic Perspectives and Policy*** 41(3):391–403.

Outreville, J-F., and E. Le Fur. 2020. “Hedonic Price Functions and Wine Price Determinants: A Review of Empirical Research.” ***Journal of Agricultural & Food Industrial Organization*** 18(2).

Rosen, S. 1974. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.” ***Journal of Political Economy*** 82(1):34–55.

Schamel, G. 2006. “Geography versus Brands in a Global Wine Market.” ***Agribusiness: An International Journal*** 22(3):363–374.

——— 2009. “Dynamic Analysis of Brand and Regional Reputation: The Case of Wine.” ***Journal of Wine Economics*** 4(1):62–80.

Schamel, G., and K. Anderson. 2003. “Wine Quality and Varietal, Regional and Winery Reputations: Hedonic Prices for Australia and New Zealand.” ***Economic Record*** 79(246):357–369.

Shapiro, C. 1983. “Premiums for High Quality Products as Returns to Reputations.” ***The Quarterly Journal of Economics*** 98(4):659–679.

Slade, P., J.D. Michler, and A. Josephson. 2019. “Foreign Geographical Indications, Consumer Preferences, and the Domestic Market for Cheese.” ***Applied Economic Perspectives and Policy***, 41(3):370–390.

Solon, G., S.J. Haider, and J.M. Wooldridge. 2015. “What are we Weighting For?” ***Journal of Human Resources*** 50(2):301–316.

Stanziani, A. 2004. “Wine Reputation and Quality Controls: The Origin of the AOCs in 19th Century France.” ***European Journal of Law and Economics*** 18:149–167.

Train, K.E. 2009. ***Discrete Choice Methods with Simulation***. Cambridge: Cambridge University Press.

Tax and Trade Bureau (TTB). 2021. “Title 27, Chapter 1.” ***Code of Federal Regulations.*** Washington, DC: Government Printing Office, 2021.

Veale, R., and P. Quester. 2008. “Consumer Sensory Evaluations of Wine Quality: The Respective Influence of Price and Country of Origin.” ***Journal of Wine Economics*** 3(1):10–29.

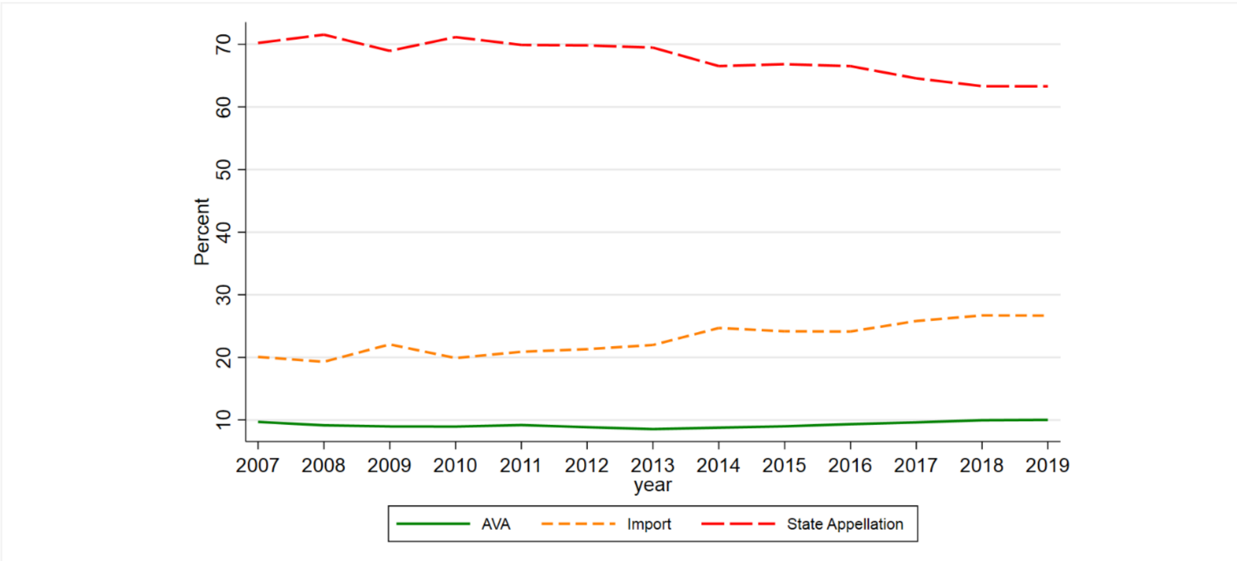
Verboven, F. 1996. “International Price Discrimination in the European Car Market.” ***The RAND Journal of Economics*** 27(2):240–268.

Villas-Boas, S.B., C. Bonnet, and J. Hilger. 2021. “Random Utility Models, Wine and Experts.” ***American Journal of Agricultural Economics*** 103(2):663–681.

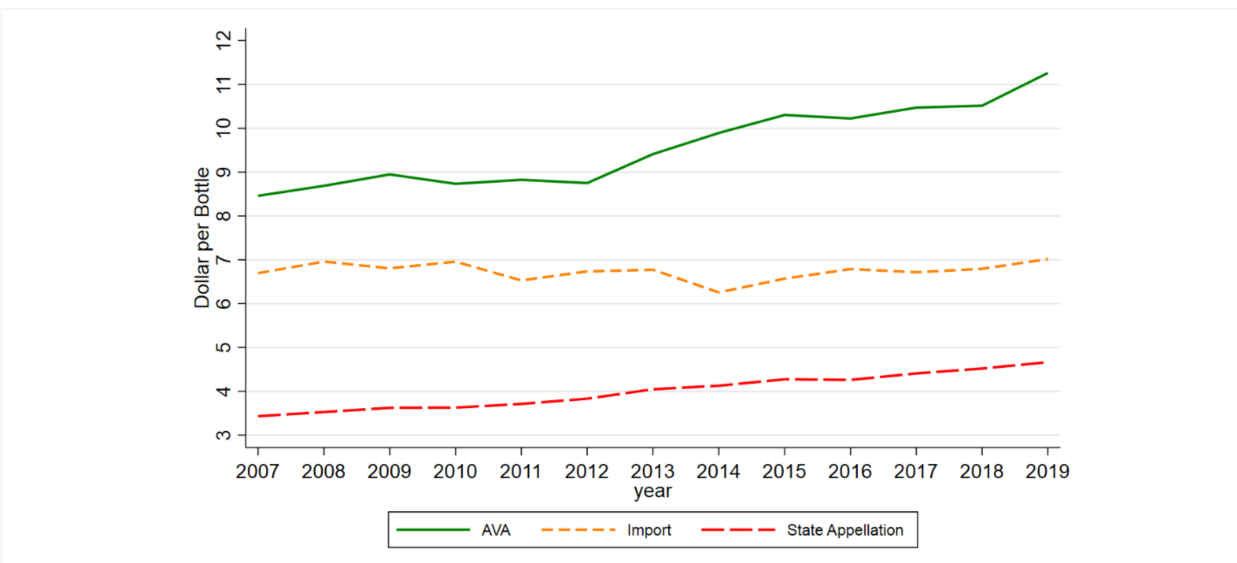
Walker, L. 2006. “Why Form an AVA?” ***Wines Vines Analytics*** October 2006.

Figure 1. Market shares and prices by geographic origin, 2007–2019.

Panel A: Market Shares



Panel B: Prices



Notes: The figures present market shares and average prices for three broad wine categories. All prices are deflated to $2019 using the Consumer Price Index.

Figure 2. Nested logit model structure.

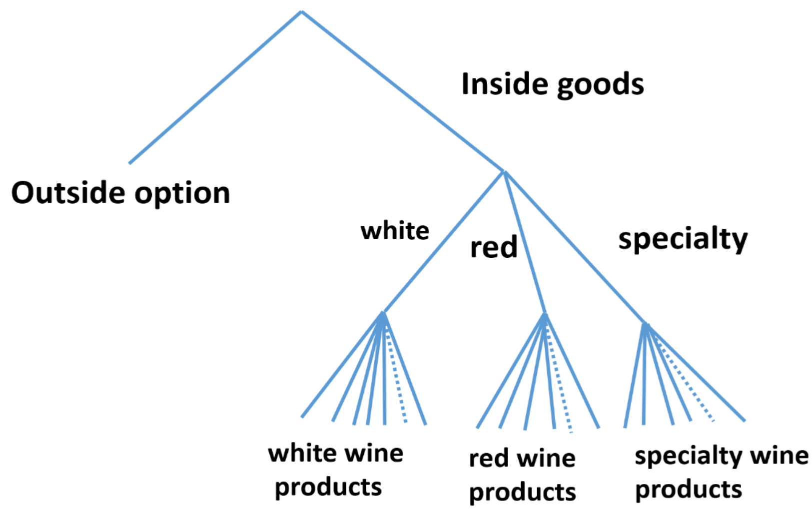


Table 1. Category Count for Characteristics Defining Products

Characteristics	Bottle	Bulk	Total
Wine type	3	3	**3**
Varietal	33	26	**33**
Packaging Size	1	1	**2**
Brand	154	20	**164**
Geographic Origin	79	48	**79**
Product Count	**2,456**	**485**	**2,941**

Notes: Table 1 reports counts of the five characteristics we use to aggregate the over 37,000 unique UPCs in the Nielsen data for our demand estimation. Varietal, brand, and geographic origin include an 'other' category for UPCs with very low market shares.

Table 2: Demand Parameter Estimates

	Nested Logit	Nested Logit	Logit	Logit-OLS
	(I)	II	(III)	(IV)
Price	-0.16***	-0.12***	-0.21***	-0.012***
	(0.0070)	(0.010)	(0.014)	(0.00059)
σ_1	0.65***	0.59***		
	(0.0060)	(0.0062)		
σ_2	0.47***	0.53***		
	(0.024)	(0.024)		
Elasticities:				
Own	-4.75	-2.91	-2.12	-0.13
Cross: Within Wine type	0.006	0.002	0.001	0.0001
Cross: Across Wine type	0.002	0.001	0.001	0.0001
Cross: Outside Good	0.0008	0.0006	0.001	0.0001
Aggregate Elasticity	-0.53	-0.38	-0.67	-0.04
Instrumental Variables	**All**	**Only BLP**	**All**	**No IVs**
Quarter FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Observations	**794,974**	**794,974**	**794,974**	**794,974**

Notes: The dependent variable is the log of the market share for every product in each market. All prices are deflated to 2019 using the Consumer Price Index. The fixed effect regressions with instrument variables are estimated using Stata's IVREGHDFE module (Correia 2018). The total number of observations in this regression is 794,985 where 11 observations are dropped as singletons. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. Marginal Utility from Second-Step Demand Estimation and WTP ($ per bottle)

	MU	SE	WTP	SE
Panel A: Packaging size, base: bottle				
bulk	-0.39***	(0.044)	-2.40***	(0.27)
Panel B: Wine type, base: Red				
Specialty	-0.35***	(0.048)	-2.16***	(0.29)
White	-0.28***	(0.050)	-1.68***	(0.31)
Panel C: Varietal Red, base: Merlot				
Cabernet Sauvignon	0.28***	(0.064)	1.71***	(0.39)
Cabernet Blend	0.080	(0.095)	0.48	(0.58)
Carmenere	0.14*	(0.078)	0.83*	(0.47)
Concord	-0.51***	(0.14)	-3.09***	(0.83)
Malbec	0.067	(0.089)	0.41	(0.54)
Moscato Red	0.0065	(0.13)	0.040	(0.79)
Petite Syrah	0.28	(0.18)	1.70	(1.10)
Pinot Noir	0.19***	(0.063)	1.17***	(0.38)
Syrah	-0.0046	(0.056)	-0.028	(0.34)
Syrah Blend	0.19	(0.26)	1.15	(1.60)
Zinfandel	0.17	(0.11)	1.01	(0.68)
Other Red	0.26***	(0.069)	1.57***	(0.42)
Other Red Imported	-0.025	(0.059)	-0.15	(0.36)
Panel D: Varietal White, base: Sauvignon Blanc				
Chardonnay	0.14***	(0.052)	0.83***	(0.32)
Chenin Blanc	-0.086	(0.12)	-0.52	(0.75)
Gewurztraminer	-0.060	(0.088)	-0.37	(0.54)
Liebfraumilch	-0.45**	(0.18)	-2.71**	(1.08)
Moscato White	0.15***	(0.056)	0.93***	(0.34)
Pinot Grigio	0.095*	(0.053)	0.58*	(0.32)
Riesling	0.038	(0.062)	0.23	(0.38)
Viognier	0.49***	(0.14)	3.00***	(0.87)
Other White	0.016	(0.062)	0.098	(0.38)
Other White Imported	0.12**	(0.058)	0.72**	(0.35)
Panel E: Varietal Specialty, base: Blush Rose				
Dessert	0.41***	(0.16)	2.51***	(0.96)
Flavored	0.063	(0.094)	0.38	(0.57)
Sangria	-0.020	(0.058)	-0.12	(0.35)
Sparkling	0.78***	(0.15)	4.77***	(0.91)
Vermouth	-0.32**	(0.16)	-1.96**	(0.96)
Other Specialty	0.18	(0.13)	1.12	(0.80)
Other Specialty Imported	0.28***	(0.092)	1.69***	(0.56)
Panel F: AVA/County, base: California				
Alexander Valley	1.14***	(0.20)	6.96***	(1.19)
Amador County	0.019	(0.14)	0.11	(0.84)
Anderson Valley	1.62***	(0.20)	9.83***	(1.21)
Arroyo Seco	0.72***	(0.22)	4.41***	(1.35)
Augusta	-0.34***	(0.11)	-2.08***	(0.70)
Carneros	1.52***	(0.24)	9.27***	(1.44)
Central Coast	0.027	(0.057)	0.16	(0.35)
Chalk Hill	2.53**	(1.28)	15.4**	(7.77)

Chalone	1.03**	(0.40)	6.27**	(2.44)
Clarksburg	0.17	(0.14)	1.03	(0.88)
Columbia Valley	0.064	(0.070)	0.39	(0.42)
Contra Costa County	-0.14	(0.27)	-0.88	(1.62)
Dry Creek Valley	1.08***	(0.24)	6.56***	(1.49)
Dunnigan Hills	-0.073	(0.18)	-0.44	(1.09)
Eagle Peak	-0.32*	(0.17)	-1.92*	(1.06)
Edna Valley	1.07***	(0.41)	6.52***	(2.50)
Eola Hills	0.31	(0.26)	1.87	(1.56)
Finger Lakes	0.37**	(0.18)	2.25**	(1.07)
Guenoc	-0.46***	(0.10)	-2.83***	(0.63)
Horse Heaven Hills	0.33*	(0.18)	2.02*	(1.12)
Knights Valley	2.02**	(0.82)	12.3**	(5.02)
Lake County	0.12	(0.13)	0.74	(0.80)
Lake Erie	-0.15	(0.13)	-0.93	(0.81)
Livermore Valley	1.19***	(0.46)	7.24***	(2.81)
Lodi	-0.095	(0.080)	-0.58	(0.48)
Mendocino	0.14	(0.092)	0.83	(0.56)
Mendocino County	-0.0092	(0.19)	-0.056	(1.16)
Monterey County	0.11*	(0.061)	0.66*	(0.37)
Napa County	0.48***	(0.18)	2.95***	(1.07)
Napa Valley	1.02***	(0.18)	6.18***	(1.07)
North Coast	0.11	(0.076)	0.65	(0.46)
Oakville	2.80***	(0.74)	17.0***	(4.51)
Old Mission Peninsula	0.33**	(0.14)	2.01**	(0.85)
Ohio River Valley	0.068	(0.22)	0.41	(1.36)
Paso Robles	0.26**	(0.11)	1.58**	(0.65)
Red Hills Lake County	0.29**	(0.14)	1.74**	(0.84)
Russian River Valley	1.08***	(0.17)	6.55***	(1.03)
Rutherford	1.03**	(0.44)	6.27**	(2.67)
Saint Lucia Highlands	0.85***	(0.23)	5.20***	(1.42)
San Antonio	-0.24	(0.17)	-1.43	(1.06)
Santa Barbara County	0.40***	(0.15)	2.43***	(0.91)
Santa Maria Valley	1.49***	(0.35)	9.03***	(2.15)
Sierra Foothills	0.42**	(0.17)	2.56**	(1.05)
Snake River Valley	-0.33**	(0.14)	-2.02**	(0.85)
Sonoma County	0.46***	(0.11)	2.78***	(0.65)
Sonoma Coast	0.89***	(0.25)	5.42***	(1.51)
Sonoma Mountain	0.37*	(0.20)	2.26*	(1.20)
Sonoma Valley	1.97***	(0.34)	12.0***	(2.05)
St Helena	0.52***	(0.16)	3.15***	(0.97)
Texas High Plains	0.093	(0.15)	0.57	(0.91)
Wahluke Slope	0.43*	(0.23)	2.60*	(1.40)
Walla Walla	0.71*	(0.41)	4.29*	(2.48)
Willamette Valley	0.79***	(0.12)	4.82***	(0.76)
Yakima Valley	0.35***	(0.10)	2.11***	(0.61)
Other AVAs	0.85***	(0.23)	5.18***	(1.40)
Panel G: State Appellation, base: California				
Florida	0.11	(0.12)	0.69	(0.73)
Indiana	0.17	(0.16)	1.05	(0.99)
Michigan	0.074	(0.15)	0.45	(0.91)

Missouri	0.16	(0.14)	0.99	(0.87)
Nebraska	0.47*	(0.28)	2.83*	(1.69)
New York	-0.31***	(0.083)	-1.87***	(0.51)
North Carolina	-0.016	(0.10)	-0.099	(0.63)
Ohio	-0.15	(0.12)	-0.89	(0.72)
Oregon	0.17	(0.11)	1.02	(0.70)
Texas	-0.14	(0.12)	-0.85	(0.73)
Washington	0.020	(0.096)	0.12	(0.58)
Other States	-0.032	(0.091)	-0.20	(0.55)
Panel H: Foreign Country, base: California				
Argentina	-0.046	(0.067)	-0.28	(0.41)
Australia	-0.14**	(0.057)	-0.85**	(0.35)
Chile	-0.23***	(0.050)	-1.43***	(0.30)
France	0.57*	(0.30)	3.49*	(1.81)
Germany	0.14	(0.095)	0.84	(0.58)
Italy	0.26***	(0.087)	1.59***	(0.53)
New Zealand	0.0057	(0.21)	0.035	(1.28)
Portugal	0.028	(0.15)	0.17	(0.94)
South Africa	-0.25***	(0.091)	-1.51***	(0.55)
Spain	-0.029	(0.097)	-0.18	(0.59)
Other Countries	-0.0045	(0.094)	-0.027	(0.57)
Constant	0.043	(0.041)	0.26	(0.25)
Observations	**2,930**		**2,930**	
Brand Fixed Effect				
R2	**0.61**		**0.61**	

Notes: The dependent variables are estimated product fixed effects. All prices are deflated to 2019 using the Consumer Price Index. These fixed effect regressions are estimated using Stata's REGHDFE command (Correia 2016). The total number of observations in this regression is 2,941, where 11 observations are dropped as singletons. Robust standard errors are in parentheses, with * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. Prices and Welfare Effects for Geographic origin of US Wines, 2007–2019

	Baseline	**Counterfactual**
AVERAGE PRICES ($/bottle)		
Imported	7.83	7.81
California	4.36	
Other state appellations	6.60	
AVAs	11.36	
United States		4.90
WELFARE CHANGE ($ billion)		
Consumer Surplus	1.18	
Variety effect	1.16	
Price effect	0.02	
Industry Revenue	3.95	
Total	5.13	

Note: Table 4 summarizes the impact of information about the geographic origin of US wines in the baseline as compared with the counterfactual. Prices and welfare metrics are expressed in 2019 dollars.

Supplementary Appendix for:

"Geographical Indications and Welfare: Evidence from the US Wine Market"

Not for Publication – To Be Made Available Online

This version: June 23, 2022

Appendix A: Descriptive Statistics

Table A1. Percent share of each module in total quantity and sales value, 2007–2019

Module name	% Quantity	% Sales	Module name	% Quantity	% Sales
Domestic Dry Table	69.36	62.23	Sweet Dessert-Domestic	0.76	0.78
Imported Dry Table	19.64	24.65	Vermouth	0.58	0.61
Sparkling	3.90	6.93	Kosher	0.29	0.33
Sangria	2.92	1.76	Sweet Dessert-Imported	0.22	0.56
Flavored/Refreshment	2.28	2.05	Aperitifs	0.018	0.07

Source: Nielsen Consumer Panel Data. Summary statistics are weighted using Nielsen projection factors. % Quantity shows the share of each module in volume (quantities), and % Sales shows the share of modules in value (revenue).

Table A2. Percent share in quantity and sales value by wine type, 2007–2019

Wine Type	% Quantity	% Sales
Red	42.92	46.49
White	33.43	32.67
Specialty	23.65	20.83

Source: Nielsen Consumer Panel Data. Summary statistics are weighted using Nielsen projection factors. % Quantity shows the share of wine type in volume (quantities), and % Sales shows the share of wine type in value (revenue).

Table A3. Percent share in quantity and sales value by packaging size, 2007–2019

Size	% Quantity	% Sales
Bottle	41.79	64.36
Bulk	58.20	35.63

Source: Nielsen Consumer Panel Data. Summary statistics are weighted using Nielsen projection factors. % Quantity shows the share of packaging size in volume (quantities), and % Sales shows the share in value (revenue).

Table A4. List of top 50 brands for bottled wine in terms of market share

Brand	Brand	Brand
14 Hands	Blackstone	Columbia Crest
19 Crimes.	Bodega Norton	Columbia Crest H3
7 Deadly Zins	Bogle	Concannon
Alamos	Bolla	Concha Y Toro Frontera
Alice White	Bonterra	Cook'S
Andre	Boone'S Farm	Covey Run
Apothic Red	C K Mondavi	Crane Lake
Arbor Mist	Castello Del Poggio	Cristalino
Ballatore	Castle Rock	Cruz Garcia
Banrock Station	Cavit	Ctl Br
Barefoot	Challis Lane	Cul-De-Sac
Bartenura	Chandon	Cupcake Vineyards
Bay Bridge Vineyards	Charles Shaw	Cutler Creek Vineyards
Beaulieu Vineyard Bv	Chateau St Jean	Da Vinci
Bella Sera	Chateau Ste Michelle	Dark Horse
Beringer	Cline	Don Simon
Black Swan	Clos Du Bois	

Note: Brand names reported in alphabetical order.

Table A5. List of top 20 brands for bulk wine in terms of market share

Brand	Brand
Almaden	Franzia
Arbor Mist	Gallo Family Vineyards
Barefoot	Liberty Creek
Beringer	Livingston Cellars
Black Box	Other_brands
Bota Box	Peter Vella
Carlo Rossi	Robert Mondavi
Concha Y Toro Frontera	Sutter Home
Corbett Canyon	Vendange
Control brand	Yellow Tail

Note: Brand names reported in alphabetical order.

Table A6. Share of varietals from US and foreign countries (bottled wine), 2007–2019

Varietal	%share	Varietal	%share	Varietal	%share
		% share in Red Wine			
United States		**Australia**		**Other Red Imported *from:***	
Cabernet Sauvignon	22.38	Syrah	2.72	France	2.69
Other Red	15.72	Cabernet Sauvignon	1.75	Germany	0.13
Merlot	14.59	Merlot	1.38	Italy	5.92
Pinot Noir	7.39	Cabernet Blend	1.28	New Zealand	0.12
Zinfandel	4.33	Other Red Imported	0.71	Other Countries	0.30
Syrah	3.32	Pinot Noir	0.54	Portugal	0.33
Petite Syrah	0.72	Syrah Blend	0.27	South Africa	0.66
Moscato Red	0.48	Malbec	0.22	Spain	3.12
Concord	0.36	**Chile**			
Cabernet Blend	0.21	Cabernet Sauvignon	1.64		
Argentina		Other Red Imported	0.62		
Malbec	3.32	Carmenere	0.60		
Other Red Imported	0.63	Merlot	0.59		
Cabernet Sauvignon	0.47	Cabernet Blend	0.24		
		Pinot Noir	0.24		
		% share in White Wine			
United States		**Australia**		**Germany**	
Chardonnay	33.45	Chardonnay	2.93	Riesling	2.45
Pinot Grigio	9.59	Pinot Grigio	1.27	Other White Imported	0.82
Sauvignon Blanc	7.22	Moscato White	1.23	Liebfraumilch	0.37
Riesling	6.26	Riesling	0.57	**New Zealand**	0.11
Other White	5.50	Other White Imported	0.30	Sauvignon Blanc	4.10
Moscato White	4.81	Sauvignon Blanc	0.21	Other White Imported	0.13
Gewurztraminer	1.45	**Chile**		**Other White Imported *from:***	
Chenin Blanc	0.89	Sauvignon Blanc	0.80	Italy	9.94
Viognier	0.20	Chardonnay	0.42	Other Countries	0.20
Argentina		Moscato White	0.25	Portugal	0.76
Other White Imported	0.73	Other White Imported	0.08	South Africa	0.52
Chardonnay	0.25	Riesling	2.45	Spain	0.85
		Other White Imported	0.80		
		% share in Specialty Wine			
United States		**Australia**		**Other Specialty Imported *from***	
Sparkling	28.40	Blush Rose	0.32	France	3.53
Blush Rose	23.78	Sparkling	0.03	Germany	0.04
Flavored	12.01	Other Specialty Imported	0.01	Italy	8.10
Sangria	3.13	Flavored	0.01	New Zealand	0.05
Vermouth	2.90	**Chile**		Other Countries	7.32
Dessert	2.85	Blush Rose	0.62	Portugal	2.64
Other Specialty	0.49	Other Specialty Imported	0.04	South Africa	0.12
Argentina				Spain	3.28
Blush Rose	0.32				

Source: Nielsen Consumer Panel Data. The % share is computed in volume (quantities) and is weighted using Nielsen projection factors.

Table A7. Share of varietals from US and foreign countries (bulk wine), 2007–2019

Varietal	% share	Varietal	% share	Varietal	% share
		% share in Red Wine			
United States		**Australia**		**Other Red Imported *from***	
Other Red	36.22	Syrah	2.91	France	0.37
Cabernet Sauvignon	19.11	Merlot	1.58	Germany	0.03
Merlot	17.65	Cabernet Sauvignon	1.22	Italy	4.45
Pinot Noir	2.33	Cabernet Blend	0.67	Other countries	0.13
Zinfandel	1.44	Other Red Imported	0.34	Portugal	0.00
Mosacto Red	0.51	Pinot Noir	0.27	South Africa	0.02
Syrah	0.51	**Chile**		Spain	2.26
Concord	0.34	Cabernet Blend	2.34		
Malbec	0.32	Merlot	1.16		
Cabernet Blend	0.28	Cabernet Sauvignon	0.80		
Argentina		Other Red Imported	0.54		
Malbec	1.11				
Other Red Imported	1.10				
		% share in White Wine			
United States		**Chile**		**Other White imported *from***	
Other White	30.88	Chardonnay	1.08	Other countries	0.14
Chardonnay	28.40	Sauvignon Blanc	0.48	Portugal	0.03
Pinot Grigio	8.92	Other White Imported	0.34	South Africa	0.03
Moscato White	5.55	**France**		Spain	0.01
Sauvignon Blanc	3.07	Other White Imported	0.07		
Riesling	0.88	**Germany**			
Argentina		Riesling	0.58		
Other White Imported	1.91	Other White Imported	0.15		
Australia		**Italy**			
Chardonnay	8.98	Other White Imported	4.66		
Pinot Grigio	1.95	**New Zealand**			
Moscato White	0.85	Sauvignon Blanc	0.40		
Other White Imported	0.44				
Riesling	0.21				
		% share in Specialty Wine			
United States		**Argentina**		**Portugal**	
Blush Rose	66.89	Blush Rose	0.72	Sangria	0.03
Sangria	12.54	**Australia**		**Other Specialty imported from**	
Flavoured	7.47	Blush Rose	0.09	France	0.11
Dessert	2.03	**Chile**		Germany	0.12
Sparkling	2.34	Blush Rose	0.13	Italy	0.39
Vermouth	0.79	**Portugal**		Other countries	4.82
Other Specialty	0.37	Blush Rose	0.04	South Africa	0.02
		Dessert	0.97	Spain	0.09

Source: Nielsen Consumer Panel Data. The % share is computed in volume (quantities) and is weighted using Nielsen projection factors.

Table A8. Percent share by geographic origin (bottled wine)

Geographic Origin	% share	Geographic Origin	% share	Geographic Origin	% share
AVA		AVA		State Appellation	
Columbia Valley	3.21	Arroyo Seco	0.05	California	46.76
Central Coast	2.43	Rutherford	0.05	New York	1.85
Napa Valley	1.73	Anderson Valley	0.04	Washington	1.37
Sonoma County	1.66	Walla Walla	0.04	North Carolina	0.39
Monterey County	1.29	Sonoma Valley	0.04	Michigan	0.36
North Coast	1.00	Mendocino County	0.04	Indiana	0.36
Lodi	0.97	Yakima Valley	0.03	Oregon	0.32
St Helena	0.70	Eagle Peak	0.03	Texas	0.30
Paso Robles	0.45	Clarksburg	0.03	Missouri	0.23
Edna Valley	0.32	Wahluke Slope	0.03	Ohio	0.17
Sonoma Coast	0.30	Knights Valley	0.03	Florida	0.15
Santa Barbara County	0.28	San Antonio	0.03	Nebraska	0.14
Russian River Valley	0.26	Snake River Valley	0.03	**Other States**	**0.56**
Mendocino	0.25	Eola Hills	0.02	Foreign Country	
Alexander Valley	0.21	Ohio River Valley	0.02	Italy	7.66
Carneros	0.21	Sonoma Mountain	0.02	Australia	6.42
Willamette Valley	0.21	Contra Costa County	0.02	Chile	2.50
Napa County	0.20	Saint Lucia Highlands	0.02	Argentina	2.49
Horse Heaven Hills	0.17	Dunnigan Hills	0.02	France	2.44
Lake County	0.11	Chalone	0.02	Spain	2.41
Livermore Valley	0.10	Augusta	0.02	New Zealand	1.44
Dry Creek Valley	0.10	Old Mission Peninsula	0.01	Germany	1.27
Guenoc	0.07	Red Hills Lake County	0.01	Portugal	0.93
Santa Maria Valley	0.06	Sierra Foothills	0.01	South Africa	0.51
Texas High Plains	0.06	Oakville	0.01	**Other Countries**	**1.66**
Finger Lakes	0.05	Chalk Hill	0.01		
Amador County	0.05	Lake Erie	0.01		
		Other AVAs	**0.16**		

Source: Nielsen Consumer Panel Data. The % share is computed in volume (quantities) and weighted using Nielsen projection factors. AVAs with less than 0.01 percent share in total bottle wine are aggregated as “Other AVAs.” Any state with less than 0.10 percent share is aggregated under the generic “Other States” designation.

Table A9. Percent share by geographic origin (bulk wine)

Geographic Origin	% share	Geographic Origin	% share	Geographic Origin	% share
AVA		State Appellation		Foreign Country	
St Helena	1.14	California	75.36	Australia	7.03
Paso Robles	0.76	New York	2.52	Italy	3.45
Monterey County	0.60	Texas	0.26	Chile	2.59
Napa Valley	0.45	Nebraska	0.10	Argentina	1.72
Santa Barbara_ County	0.13	Washington	0.07	Spain	0.92
Lodi	0.12	Ohio	0.04	Portugal	0.29
Columbia Valley	0.11	Missouri	0.01	Germany	0.29
Central Coast	0.06	Michigan	0.01	France	0.20
Sonoma County	0.03	Oregon	0.01	New Zealand	0.14
Finger Lakes	0.02	Other States	0.11	South Africa	0.02
Napa County	0.02			Other Countries	1.37
Snake River Valley	0.01				

Source: Nielsen Consumer Panel Data. The % share is computed in volume (quantities) and weighted using Nielsen projection factors.

Table A10. Percent share of total imports by foreign country and packaging size

Country	% Total Import	% Share in Bottle	% Share in Bulk
Australia	29.54	11.69	17.84
Italy	22.74	13.97	8.77
Chile	11.14	4.55	6.58
Argentina	8.88	4.53	4.35
Spain	6.74	4.250	2.34
France	4.95	4.44	0.50
Germany	3.05	2.31	0.74
New Zealand	2.98	2.63	0.34
Portugal	2.44	1.69	0.74
South Africa	0.98	0.92	0.05
Other countries	6.51	3.03	3.48

Source: Nielsen Consumer Panel Data. The % share is computed in volume (quantities) and weighted using Nielsen projection factors.

Appendix B: Robustness Checks

Table B1. Demand parameter estimates sensitivity to market size

Market Size	Nested Logit baseline	Nested Logit double	Nested Logit triple
	(I)	(II)	(III)
Price	-0.16***	-0.14***	-0.12***
	(0.0070)	(0.0057)	(0.0049)
σ_1	0.65***	0.71***	0.75***
	(0.0060)	(0.0049)	(0.0042)
σ_2	0.47***	0.56***	0.62***
	(0.024)	(0.020)	(0.017)
Elasticities:			
Own	-4.75	-4.75	-4.74
Cross: Within Wine type	0.006	0.010	0.010
Cross: Across Wine type	0.002	0.003	0.003
Cross: Outside Good	0.0008	0.0005	0.0003
Aggregate Elasticity	-0.53	-0.51	-0.51
Instrumental Variables	**All**	**All**	**All**
Quarter FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Observations	**794,974**	**794,974**	**794,974**

Note: The dependent variable is the log of the market share for every product. All prices are deflated to 2019 using the Consumer Price Index. The fixed effect regressions with instrument variables are estimated using Stata's IVREGHDFE module (Correia 2018). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2. Demand parameter estimates with year trends

	Nested Logit	**Nested Logit**	**Logit**	**Logit-Ols**
	(I)	(II)	(II)	(III)
Price	-0.16***	-0.20***	-0.061***	-0.012***
	(0.0066)	(0.0098)	(0.014)	(0.00058)
σ_1	0.80***	0.67***		
	(0.0054)	(0.0065)		
σ_2	0.63***	0.47***		
	(0.021)	(0.024)		
Year trend	-0.044***	-0.046***	-0.019***	-0.014***
	(0.00077)	(0.0011)	(0.0016)	(0.00048)
Elasticities:				
Own	-8.13	-6.16	-0.62	-0.12
Cross: Within Wine type	0.014	0.009	0.0003	0.0001
Cross: Across Wine type	0.004	0.003	0.0003	0.0001
Cross: Outside Good	0.0008	0.001	0.0003	0.0001
Aggregate Elasticity	-0.52	-0.62	-0.19	-0.03
Instrumental Variables	**All**	**Only BLP**	**All**	**No IVs**
Quarter FE	Yes	Yes	Yes	Yes
Year FE	No	No	No	No
Region FE	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Observations	**794,974**	**794,974**	**794,974**	**794,974**

Note: The dependent variable is the log of the market share for every product. All prices are deflated to 2019 using the Consumer Price Index. The fixed effect regressions with instrument variables are estimated using Stata's IVREGHDFE module (Correia 2018). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3. Nested logit first stage estimates

	Nested Logit (I)			Nested Logit (II)		
	p_{jm}	s_{jm} / S_{h1m}	S_{h1m} / S_{1m}	p_{jm}	s_{jm} / S_{h1m}	S_{h1m} / S_{1m}
Product count	-0.0002	-0.002	-0.001	0.0008	-0.002	-0.0008
	(0.0003)	(0.0002)	(0.0000)	(.0001)	(.0001)	(.00012)
Product count by wine type	-0.0011	-.004	0.0029	-0.0010	-0.004	0.002
	(0.0003)	(0.0002)	(0.0000)	(0.0003)	(0.00002)	(0.00002)
Product count by packaging size	-0.002	0.0001	0.0003	-0.002	0.0001	-0.0009
	(0.0004)	(0.0001)	(0.0000)	(0.0002)	(.00001)	(.00001)
Product count by brand	0.002	0.00057	-0.0001	0.001	0.001	-0.0001
	(0.0008)	(0.0002)	(0.000)	(0.0005)	(0.0002)	(0.0002)
Product count by geographic origin	0.001	0.004	-0.0002	0.002	0.002	-0.0001
	(0.0004)	(0.0002)	(0.0000)	(0.0004)	(0.00006)	(0.00006)
Product count by wine type and packaging size	0.001	0.001	-0.0009	0.001	0.001	0.0019
	(0.0004)	(0.0003)	(0.00000)	(0.0002)	(0.0002)	(0.0002)
Product count by wine type and varietal	-0.001	0.006	0.0003	-0.001	0.007	0.0003
	(0.00057)	(0.0003)	(.00000)	(0.0005)	(0.00029)	(0.00029)
Product count by wine type and brand	0.004	0.007	0.0004	0.004	0.007	0.0004
	(0.002)	(0.0005)	(.00010)	(0.0011)	(0.0004)	(0.0004)
Product count by wine type and geographic origin	-0.002	0.0008	0.0005	-0.002	0.0008	0.0005
	(0.00092)	(0.0003)	(0.0001)	(0.0003)	(0.0005)	(0.0005)
Distribution Cost	0.00075	-.0006	0.0000			
	(0.00005)	(0.00002)	(0.00000)			
Distance traveled by competing products	0.0000	0.0000	0.0000			
	(0.000)	(0.000)	(0.000)			
Retail Density (area adjusted)	0.0059	0.003	-0.0011			
	(0.0009)	(0.0004)	(0.0000)			
Retail Density (population adjusted)	0.0000	0.0000	0.0000			
	(0.000)	(0.000)	(0.000)			
Excise tax	0.15	-0.067	-0.035			
	(0.022)	(0.014)	(0.003)			
State Control	0.65	-0.07	-0.015			
	(0.034)	(0.014)	(0.0038)			
F test of excluded instruments	**110.77**	**552.76**	**613.17**	**117.49**	**810.02**	**987.03**
Sanderson-Windmeijer F test of excluded instruments	**109.72**	**553.09**	**424.70**	**140.59**	**730.78**	**754.04**

Note: The table reports first stage estimates demand model reported in Table 2, columns I and II. The total number of observations is 794,974. Robust standard errors are included in parenthesis.

Appendix C: Data Extraction and Cleaning

We extract wine attribute information, such as type, varietal, and geographic origin, from the UPC description in the Nielsen data. Table C1 presents some examples of UPC and brand descriptions from the Nielsen data. For most UPC descriptions, the first word is generally the brand name. The remaining letters stand for country name, region, and varietal. For instance, CHL, IT, FR, and ARG are abbreviations for Chile, Italy, France, and Argentina. CHRD is Chardonnay, P-GR is Pinot Grigio, SHZ CB is Shiraz Cabernet, MLBC is Malbec, WT is White, IDT is Imported Dry Table, and DDT is Domestic Dry Table.

In some cases, we found multiple abbreviations conveying the same information, e.g., Moscato is represented in various entries as MSC, MSCT, and MUSCATO. To ensure that the abbreviations mean what we are inferring, we cross-checked them by matching the brand name and UPC description with outside sources. The site www.wine-searcher.com, in particular, has a comprehensive list of wines with brand names, grape blends, and geographic origins. Along with UPC, we also used Nielsen's 'style_desc' variable to extract the geographic origin information. For entries where we could not find geographic origin information from the Nielsen data, we manually searched for the information using www.wine-searcher.com and https://www.vivino.com/US/en/.

Wine brand name comes with small variations, and Nielsen assigns distinct codes to brands with even the slightest variation in brand description. In this study, we treat different brand variants as one. For instance, we consider Francis Coppola dmnd cltn, Francis Coppola diamond cllctn, Francis coppola presents, and Francis Coppola Director's cut as one brand. Similarly, the brand Robert Mondavi includes Robert Mondavi, Woodbridge Rbrt Mndv, Robert Mondavi Private Selection, La Famiglia di Robert Mondavi, Woodbridge Rbrt Mndv Slt Vy Sr, and Woodbridge by Robert Mondavi. Reconciling alternative variants for the same brand name reduced the number of individual brands by around 250.

Table C1 below reports several examples of wine items in the data. The first three columns display the raw data as available in the dataset. The last column, "geographic origin," is imputed by us based on the available information.

Table C1. Examples of UPC descriptions form Nielsen Data

style_descr	brand_descr	upc_descr	"geographi origin"
A-V	Geyser Peak	GYSR PK CHRD AV V WT DDT	Alexander_Valley
A-V RS	Louis M Martini	LMM CB-S AV RS V RED DDT	Alexander_Valley
A-V CA SC RS	Clos Du Bois	CDB CB-S AV CA SC RS V RED DDT	Alexander_Valley
ANDERSON VALLEY	Goldeneye	GLDE TD P-N AD-V V RED DDT	Anderson_Valley
ANDERSON VALLEY MENDOCINO	Brutocao	BRUTOCAO P-N AVM CLS V RED DDT	Anderson_Valley
C-V WASHINGTON STATE	Wild Meadows	WLD MDW MRLT C-VWS V RED DDT	Columbia_Valley
C-V	Guardian	GUARDIAN GM-CSMF C-V V RED DDT	Columbia_Valley
NAPA VALLEY	Robert Mondavi	R-M CB-S NV V RED DDT	Napa_Valley
-	Ctl Br	CTL BR CB-S NV V RED DDT	Napa_Valley
CA NAPA VALLEY OAKVILLE	Opus One	OPUS ONE RED C-NV-O G RED DDT	Oakville
LODI	Noble Vines	NL-V 337 CB-S LDI V RED DDT	Lodi
CA CENTRAL COAST	Macmurray	MCMRY P-N CA CC V RED DDT	Central_Coast
NAPA COUNTY	Crane Lake	CRNE LKE RES N-C V WT DDT	Napa_County
SC	Landing Place	LDG PL CHRD SC V WT DDT	Sonoma
SBC	Seaglass	SEAGLASS SV-B SBC V WT DDT	Santa_Barbara_County
SONOMA COAST	Bearboat	BEARBOAT P-N SCS V RED DDT	Sonoma_Coast
S-V	St. Francis	ST FRN MRLT S-V V RED DDT	Sonoma_Valley
-	Tualatin Vineyards	TLN VNYD MSCT WV V WT DDT	Willamette_Valley
MTRY COUNTY SC	Francis Coppola Director'S	FCDT TWZ MRLT M-C SC V RED DDT	Monterey_County
MTRY	Ranch 32	RANCH 32 CB-S MTRY V RED DDT	Monterey_County
MONTEREY COUNTY	It'S A Headsnapper	ITS A-HDSR P-N SCS V RED DDT	Monterey_County
MTRY RS	Fog Head	FOG HEAD P-N MTRY RS V RED DDT	Monterey_County
MTRY COUNTY	Baron Herzog	BRN-H RES M-C V WT DDT	Monterey_County
YAKIMA VALLEY	Pacific Rim	P-RIM GRNR-V YV V WT DDT	Yakima_Valley
CA	Beringer	BGR WT ZN V BLS DDT	California
CA	Menage A Trois	MNG-AT RED CA G RED DDT	California
CA	Sutter Home	S-HM WT ZN CA V BLS DDT	California
CA VINTNERS RS	Kendall-Jackson	K-J SV-B CA V-R V WT DDT	California
WASHINGTON	Columbia Crest	CL-C 2VN RSE WSH G BLS DDT	Washington
WASHINGTON STATE	14 Hands	14 HNDS P-GR WS V WT DDT	Washington
-	Taylor'S	TYLR PORT NY SD D	New York
OREGON	Wine By Joe	WN-JOE CHRD OG V WT DDT	Oregon
PENNSYLVANIA	Franklin Hill Vineyards	FHVY F-G PNSY V RED DDT	Pennsylvania
-	Horton	HRTN PORT VI SD D	Virginia
-	Sunshine Bay	SUNSHINE BAY NZ SV-B WT IDT	New Zealand
-	Il Carnevale Di Venezia	ICDV IT GR-P-GR WT IDT	Italy
-	La Vieille Ferme	LVF FR CTE DU VNTX RG RED IDT	France
-	Monmousseaux	MONMOUSSEAUX FR VVRY WT IDT	France
MENDOZA	Black Box	BKB ARG MLBC MNDZ V RED BB IDT	Argentina