Wine Retail Price Dispersion in the United States

David A. Jaeger and Karl Storchmann*

' Similar to other markets in which deviations from Jevons' "law of one price" is the norm rather than the exception, the retail wine market in the United States is characterized by enormous price dispersions. For instance, we observe retail prices for *2005 Chateau Latour* ranging from $695 in a Petaluma, California, wine store to $2000 in a wine store in Champaign, Illinois. Similarly, at the lower end of the price distribution, the observed retail price of *2007 Yellowtail Merlot* ranges from $4.99 in Buffalo, New York to $9.99 in Jersey City, New Jersey. Price dispersion on the wine market can be caused by various factors such as differences in production and distribution cost, differences in price elasticities of demand, or different market regulations and structures.

Since the ratification of the 21st Amendment repealing Prohibition, the U.S. wine market has been primarily regulated at the state level, more or less impairing or effectively abolishing competition between wine retail outlets. For example, New Jersey allows its municipalities to issue one wine store license for every 7,500 of its population. In addition to the federal wine tax, wine is levied by state-specific wine and sales taxes. Some states maintain a monopoly over the wholesale and retail of wine (18 control states); others restrict the sales of wine to certain outlets and/or certain times or even prohibit the payment for wine purchases with credit cards. Many states prohibit direct wine shipments from out-of-state producers and retailers, while others even prohibit in-state producers and retailers to ship wine to consumers. Price differences between

* Jaeger: Center for Macroeconomic Research, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany (jaegerd@uni-koeln.de) and CUNY Graduate Center. Storchmann: Economics Department, New York University, 19 W. 4th St., 6FL, New York, NY 10012 (karl.storchmann@nyu.edu).

states or counties are thus not surprising. In this paper we examine whether state- or county-specific effects fully explain the observed price dispersion or if price variations remain even after controlling for location differences. If so, is the degree of price dispersion identical for all wines or are there differences?

A large body of information-theoretic literature suggests that markets even for standardized products may exhibit considerable price dispersion. Following George J. Stigler's (1961) paper, several authors model how equilibrium price dispersion can arise as a result of heterogenous information (e.g., Steven C. Salop and George E. Stiglitz, 1977, 1982; Jennifer F. Reinganum, 1979; Hal R. Varian, 1980; Kenneth Burdett and Kenneth Judd, 1983; John Carlson and Preston McAfee, 1983; and Dale O. Stahl II, 1989). In general, price dispersion can persist in equilibrium if obtaining information is costly (through, for example, search costs) and some fraction of consumers chooses to be uninformed.

A variety of empirical studies have explicitly examined the association between consumer search and price dispersion for homogenous goods. John W. Pratt et al. (1979) examine price dispersions for 39 consumer goods in the Boston area and report coefficients of variation (CV) for the product prices between 4 and 71 percent. They also find that the price dispersion substantially increases with the average price of the good, suggesting that the search cost for expensive items is higher. This may be explained by the fact that expensive products are purchased less frequently, reducing the incentive of a buyer to search. Bev Dahlby and Douglas West (1986) find a CV of 18% for auto insurance policies in Alberta. After ruling out quality or cost differences they conclude that this price dispersion is almost exclusively due to costly consumer search. Alan T. Sorenson (2000) examines the retail prices of pharmacies in two geographically distinct markets and finds a CV of 22%. While at most one third of the observed price dispersion is due to pharmacy heterogeneity most is due to costly search. Sorenson also finds that frequently purchased prescriptions exhibit lower price variations.

Most relevant for our study, it has been hypothesized that the Internet and the emergence of online markets substantially lower search cost resulting in lower price dispersion (e.g., Yannis Bakos, 1997). Xing Pan et al. (2002) analyze the price dispersion of 581 goods in eight product categories in online markets. After controlling for sellers' heterogeneity and especially service quality, however, they find online price dispersion to be substantial and persistent. Erik K. Clemons et al. (2002) report similar results for the market of airline tickets sold by online travel agents. Kathy Baylis and Jeffrey M. Perloff (2002) analyze Internet prices of a specific type of digital camera and a flatbed scanner over a 14-week period and also find significant price dispersion, which even increases when controlling for service quality. In contrast to Varian's (1980) model of mixed strategies, they find a pure-strategies equilibrium, with high-priced firms and low-price firms remaining fixed in the overall ranking over time. They conclude that information costs (the time taken to negotiate the website to discover stock and, to some extent, price information) are an important determinant of online price dispersion, and that firms may discriminate among consumers based on their knowledge, search costs, or patience.

Since a high degree of price dispersion indicates large potential gains to search by consumers it may also suggest that the market in question is inefficient with regard to information. Empirical research has shown that consumer search in most cases stops before full information is obtained; sometimes no search takes place at all (Brian T. Ratchford, 2009). Given that search is costly, however, the optimum search point is reached when marginal search cost equals its marginal benefit. Ratchford and Narasimhan Srinivasan (1993), Edward J. Fox and Stephen J. Hoch (2005) and Dinesh K. Gauri et al. (2007) provide empirical evidence that is consistent with this normative rule.

In this analysis, we draw on a large database of wine retail prices from 2006 to 2008 provided by *wine-searcher.com* to examine the relationship between price and price dispersion. We first examine the role of local characteristics (the number of retail wine establishments, per capita income, and local demographics) on wine prices; we control for the regulatory environment by using state fixed effects. First, we explore the relationship between prices and local market characteristics such as per capita income, age and race, as wine stores per population and state regulations (captured by state fixed effects). After also controlling for wine-vintage fixed effects, we then examine whether the residual variation in prices is related to the wine's average price. In general, we find a significant and positive relationship between residual variation in prices and (adjusted) price levels.

Given that the search cost is essentially fixed per wine and independent of its price (e.g., searching a website), it is possible that search is more profitable for expensive wines, resulting in smaller price dispersion with increasing average prices. On the other hand, less expensive wines face a stiffer competition from close substitutes than expensive wines do. In contrast to a $200 wine, when a consumer shops for a $5 wine the brand and vintage is likely to be of less importance. Monopoly pricing power may therefore increase with price, potentially leading to a price dispersion that increases with a wine's average price. Alternatively, learning through experience may play a role and lead to the same dispersion-price relationship. Low-price wines sell at much higher quantities than high-end wines. Information about quality and prices of lower-tier wines may more easily penetrate the market (for learning from buying and word-to-mouth, see Ratchford, 2009).

I. Data and Descriptive Statistics

We use wine retail prices from 2006 to 2008 provided by *wine-searcher.com*, an Internet wine price search site on which wine retail outlets worldwide can post prices of their wines. In the U.S., *wine-searcher.com* currently lists approximately 2.5 million prices posted by about 6,300 wine stores. Since many wines are available only in a few stores we restrict our analysis to 186 wine brands of various vintages. For all but one of these wines we observe well over 200 prices, and for many we observe more than 1,000 prices.[1] Overall, our sample contains approximately 106,000 prices on red and white wines. We report some basic descriptive statistics on prices levels and price dispersion in Table 1. Most of the wines in our sample are produced in the US and two-thirds of them are red. We observe substantial differences in price dispersion, measured by the coefficient of variation. Compared to the results of other empirical analyses, the overall price dispersion of 23.4% is rather high. It is higher for red than for white wines and higher for French wines compared to domestic wines and other imports (mainly from Australia and Italy). Also, expensive wines exhibit higher price dispersion than do wines in lower price brackets suggesting the dominance of the substitution effect or learning from buying over information hypotheses.

[1] The data contain observations for sizes in addition to the standard 750ml bottle. We have dropped all observations for non-standard sizes. In addition, for each wine we have dropped the 5% lowest and 5% highest observed prices, to be sure that we were not capturing (mislabeled) case prices or other measurement issues. We have also dropped any observations in which the description indicated that the bottle was damaged or irregular in any way. For wines with both vintage and non-vintage prices reported, we dropped any non-vintage prices when these constituted less than 25% of the total number of observations for that wine. We also eliminated from the data rosé, sparkling, and fortified wines.

II. Determinants of Wine Prices

To examine how local market characteristics affect wine prices, we estimate the simple equation

$$\log(p_{ivcsy}) = \beta_0 + \beta_1 E_{cy} + \beta_2 I_c + \beta_3 W_c + \beta_4 A_c + \beta_5 O_{ivy} + \beta_6 NV_i + \theta_y + \delta_s + \lambda_{iv} + \varepsilon_{ivcsy} \quad (1)$$

where i indicates wine, v indicates vintage, c indicates county, s indicates state, and y indicates year of price posting. The variable E is the number of retail wine establishments in county c in year y divided by the county population in 2000, taken from the county business patterns data of the U.S. Census Bureau. I is per capita income in the county in 2000, W is the white share of the population in the county in 2000, A is the share of the county population in 2000 that is 25 or over (the population most likely to drink wine), O is how old the wine is in year y (non-vintage wines are coded to zero), NV is an indicator for non-vintage wines, θ is a year fixed effect, δ is a state fixed effect (capturing differences in state regulations), λ is a wine × vintage fixed effect, and ε is the idiosyncratic term. Note that in some specifications, we will use only simple wine fixed effects, without letting the coefficient vary across vintages.

The results of estimating variants of equation (1) are presented in Table 2. The first three columns contain results for red wine, while the last three contain results for white wine. In columns (1) and (4) we constrain the λ's to be equal to zero, as well as the coefficients wine age and non-vintage. It is clear from both columns that prices vary with local market conditions, even despite our sample being drawn from sellers who list their prices on the Internet. Local market conditions only explain between 7 (red) and 13 (white) percent of the variation in prices, however. This is not surprising – in this regression we are treating all wines the same, regardless of where or by whom they were produced.

In columns (2) and (5) we add to the analysis fixed effects for each wine, but constrain these to be equal across vintages. The model now accounts for 95 percent of the variation in log prices – clearly the majority of variation in wine prices comes from differences in origin and quality. The coefficients on income, white share, and age change somewhat, suggesting that different wines are sold in different locations. In columns (3) and (6), we allow for a full set of wine × vintage interactions. The results are qualitatively similar to those in column (2)

III. The Relationship between Price and Variance

Our fundamental research question is whether there is a relationship between residual price variation and price level. At base, this is a test of heteroskedasticity. As we have seen above, we can reject the null hypothesis of homoskedasticity in all of the regressions in Table 2. In Table 3, we report the slope coefficients from a regression of the average squared residual on the fixed effect for each wine × vintage combination, both taken from columns (3) and (6) of Table 2 for red and white wines, respectively. That is, we are estimating

$$\bar{e}_{iv}^2 = \phi + \varphi\hat{\lambda}_{iv} + \xi_{iv} \tag{2}$$

where e are the residuals from the estimation of equation (1). Here we find that, overall, there is a positive relationship between residual variation in prices and their level. To put the magnitude of the coefficient in context, the average value of the dependent variable for red wines (that is, the average mean squared residual) is .0287. Thus, the estimate coefficient on the full sample of 1117 wine × vintage combinations is about one-tenth of his average. For white wines, the average mean squared residual is .0209 and the estimated coefficient is for the full sample is about three-tenths of this (.0064). For both red wines and white wines, we find a stronger statistical relationship between dispersion and average price for vintage wines than for non-vintage wines. In Figures 1 and 2 we also present the results of non-parametric regressions for red and white wines, respectively, using a bandwidth in each of .75. These both show a very slight positive relationship between average price and price dispersion.

V. Conclusion

In this paper we show that there is a fair amount of price dispersion for red and white wine in the U.S., with an average per-wine coefficient of variation of 23 percent. Some of this is due to differential market conditions. But our evidence suggests that dispersion also depends (weakly) on price levels, after controlling for consumer, market, and state heterogeneity. These results are consistent with the theory of "learning-by-buying" in which goods that are purchased more often are predicted to have less price heterogeneity. The results are less consistent with a search costs story. To be consistent with our results, search costs would have to be higher for expensive wines relative to less-expensive wines. Given that the search mechanisms are the same for both inexpensive and expensive wines, this seems to us to be less plausible.

References

Bakos, Yannis. 1997. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces." *Management Science,* 43(12): 1676-1692.

Baylis, Kathy and Jeffrey M. Perloff. 2002. "Price Dispersion on the Internet: Good Firms and Bad Firms." *Review of Industrial Organization*, 21(3):305-324.

Burdett, Kenneth and Kenneth Judd. 1983. "Equilibrium Price Dispersion." *Econometrica*, 51(4): 955-969.

Carlson, John and Preston McAfee. 1983. "Discrete Equilibrium Price Dispersion." *Journal of Political Economy*, 91(3): 480-493.

Clemons, Erik K., Il-Horn Hann, and Lorin M. Hitt. 2002. "Price Dispersion and Differentiation in Online Travel: An Empirical Investigation." *Management Science*, 48(4): 534-549.

Dahlby, Bev and Douglas West. 1986. “Price Dispersion in an Automobile Insurance Market.” *Journal of Political Economy*, 94(2): 418-438.

Fox, Edward J. and Stephen J. Hoch. 2005. “Cherry-Picking.” *Journal of Marketing*, 69(1): 46-62.

Gauri, Dinesh K., K Sudhir, and Debabrata Talukdar. 2007. “The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data.” *Journal of Marketing Research*, 45(2): 226-240.

Pan, Xing, Brian T. Ratchford and Venkatesh Shankar. 2002. “Can Price Dispersion in Online Markets be Explained by Differences in E-Tailer Service Quality?” *Journal of the Academy of Marketing Science*, 30(4): 433-455.

Pratt, John W., David A. Wise, and Richard J. Zeckhauser. 1979. "Price Differences in Almost Competitive Markets." *Quarterly Journal of Economics*, 93(2): 189-211.

Ratchford Brian T. and Narasimhan Srinivasan. 1993. “An Empirical Investigation of Returns to Search.” *Marketing Science,* 12(1): 73-87.

Ratchford, Brian T. 2009. “Consumer Search and Pricing” in *Handbook of Pricing Research in Marketing,* Rao, Vithala R. editor. Edward Elgar: Cheltenham; 91-107.

Reinganum, Jennifer F. 1979. “A Simple Model of Equilibrium Price Dispersion.” *Journal of Political Economy*, 87(4): 851-858.

Salop, Steven C. and Joseph E. Stiglitz. 1977. “Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion.” *Review of Economic Studies*, 44(3): 493-451.

Salop, Steven C. and Joseph E. Stiglitz. 1982. "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents." *American Economic Review*, 72(5): 1121-1130.

Sorensen, Alan T. 2000. “Equilibrium Price Dispersion in Retail Markets for Prescription Drugs.” *Journal of Political Economy*, 108(4): 833-850.

Stahl, Dale O. II. 1989. "Oligopolistic Pricing with Sequential Consumer Search." *American Economic Review*, 79(4): 700-712.

Stigler, George J. 1961. “The Economics of Information." *Journal of Political Economy*, 69(3): 213-225.

Varian, Hal R. 1980. “A Model of Sales.” *American Economic Review*, 70(4): 651-659.

Figure 1: Red Wines

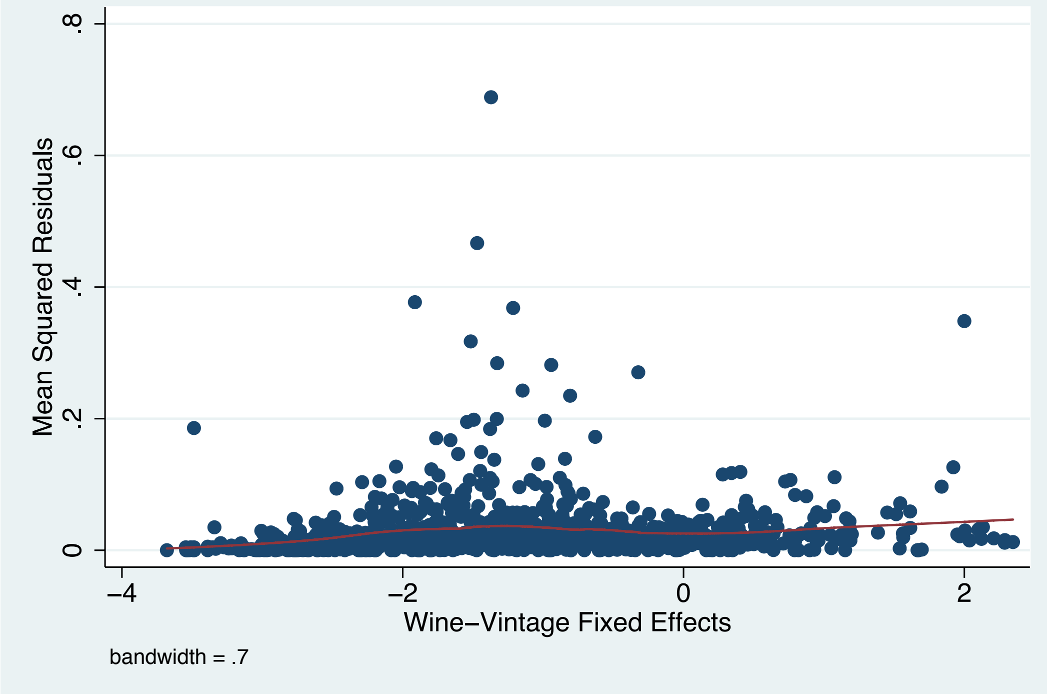

Figure 2: White Wines

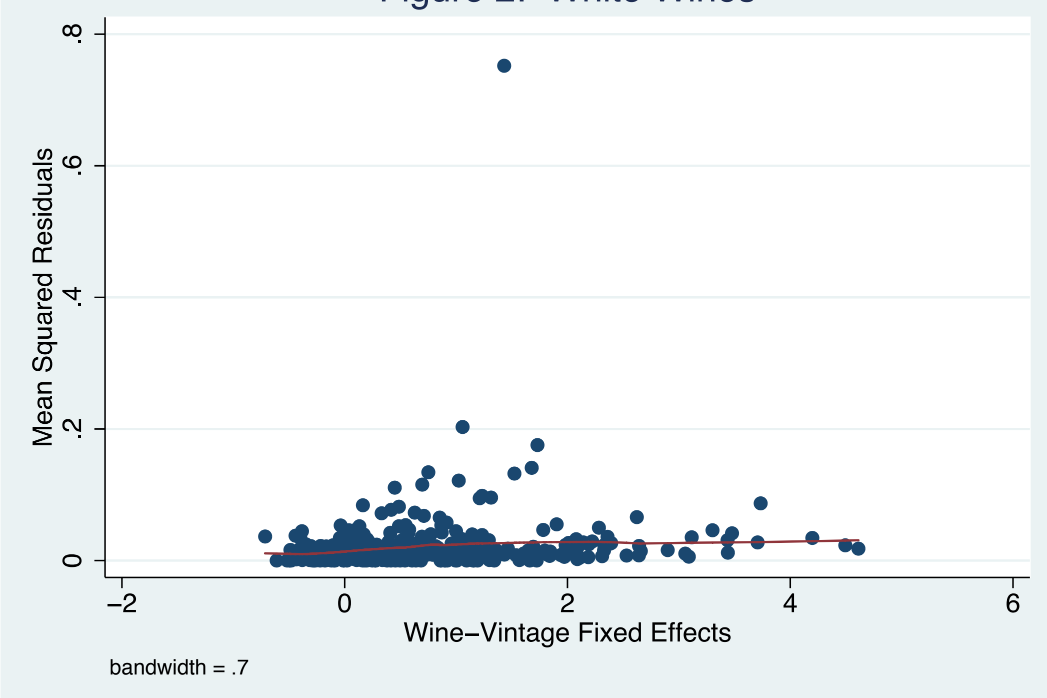

Table 1
Descriptive Statistics on Prices

Sample	Ave. Mean Price	Ave. Coeff. of Variation	Ave. *N* per Wine	Number of Wines
Full Sample				
Red	$ 80.25	0.2495	984.45	136
White	$ 27.84	0.1925	746.32	50
U.S.				
Red	$ 45.61	0.2002	914.41	66
White	$ 15.17	0.1814	746.84	31
France				
Red	$148.64	0.3492	1173.49	47
White	$130.53	0.3732	689.33	6
Other				
Red	$ 39.87	0.1874	799.13	23
White	$ 10.66	0.1358	771.38	13
Averge Price < $15				
Red	$ 8.50	0.1668	661.92	36
White	$ 8.71	0.1668	683.97	29
$15 <= Average Price < $50				
Red	$ 28.34	0.2576	842.47	32
White	$ 22.83	0.1665	883.69	16
Average Price >= $50				
Red	$142.65	0.2895	1222.02	68
White	$154.83	0.4099	668.40	5

Table 2
Determinants of Wine Prices

Variable	(1)	(2)	(3)	(4)	(5)	(6)
		Red			**White**	
Number of Wine Retailers per 2000 County Population	0.0031 (0.0002)	0.0002 (0.0001)	0.0002 (>0.0001)	0.0023 (0.0004)	-0.0002 (0.0001)	-0.0001 (0.0001)
Log Per Capita Income (2000) in County	0.1568 (0.0011)	-0.0300 (0.0068)	-0.0083 (0.0007)	0.1423 (0.0386)	-0.0208 (0.0099)	-0.0107 (0.0079)
White Share of County Population (2000)	-0.0243 (0.0073)	-0.0017 (0.0085)	-0.0189 (0.0060)	-0.2334 (0.0515)	0.0214 (0.0129)	-0.0066 (0.0101)
Share of County Population that is 25 or older (2000)	0.0247 (0.0249)	0.1714 (0.0393)	0.1987 (0.0282)	1.4587 (0.2307)	0.2767 (0.0608)	0.2412 (0.0479)
Wine Age (NV=0)		0.0075 (0.0008)	0.1411 (0.0034)		0.0231 (0.0023)	0.0533 (0.0042)
Non-Vintage		-0.0573 (0.0042)			-0.0152 (0.0067)	
Year Fixed Effects	X	X	X	X	X	X
State Fixed Effects	X	X	X	X	X	X
Wine Fixed Efffects		X			X	
Wine x Vintage Fixed Effects			X			X
R^2	0.07	0.95	0.98	0.13	0.94	0.97
N		82,698			23,919	

Source: Authors' calculations using data from wine-tracker.com Observations from from 2006-2008 measuring prices of non-vintage and vintage wines from 1998-2007.
Note: Dependent variable is log price. Estiamted via OLS. Heteroskedasticity-consistent standard errors are in parentheses.

Table 3
Relationship between Price Dispersion and Average Price

	Red		White	
Sample	**Coeff.**	***N***	**Coeff.**	***N***
All	0.0024 (0.0010)	1,117	0.0064 (0.0020)	398
Vintage	0.0025 (0.0010)	1,056	0.0062 (0.0020)	352
Non-Vintage	0.0291 (0.0151)	61	0.0171 (0.0133)	46

Source: Authors calculations of data from wine-tracker.com.
Note: Each coefficient comes from a separate regression. We regress the average squared residuals for each wine x vintage combination on the fixed effects for those combinations. These are taken from columns (2) and (6) of Table 2. Heteroskedasticity-consistent standard errors in parentheses