WORKING PAPERS

N° TSE –834

Revised - May 2019

"Reduced Form Evidence on Belief Updating under Asymmetric Information - The Case of Wine Expert Opinions"

Céline Bonnet, James Hilger and Sofia B. Villas-Boas

Toulouse School of Economics

Reduced Form Evidence on Belief Updating under Asymmetric Information - The Case of Wine Expert Opinions

Celine Bonnet, James Hilger, and Sofia B. Villas-Boas*

May 10, 2019

Abstract

We estimate the effect of revealing expert opinion labels on wine product purchases through a field experiment where a random subset of wine products within the consumers' retail shelf choice set are labeled in the treatment store. We use a detailed weekly product level panel scanner data set for labeled and unlabeled wines in the treatment and comparable control stores before and after the implementation of a product-level labeling field experiment. We combine the scanner data with additional information on the characteristics of each product, such as brand, varietal, region of production, and price to estimate the average and heterogeneous effects of the field experiment on wine consumption. Consistent with earlier work, we find there to be a positive and significant overall average effect and that demand increases more for higher score wines than for lower score wines. We advance the literature with the following: one, higher scores matter more for prices in the lower quartile of the overall wine price distribution whereas demand does not move for the higher priced wine quartile, once quality is revealed; the result is consistent with pre treatment consumer behavior where consumers infer high quality for high prices. Two, we find positive spillover effects of this experimental treatment within brand for untreated wines as the displayed average score of the wine brand increases. However, we also obtain negative spillover effects for untreated wines that belong to intensively treated brands.

Keywords: Field experiment, Labels, information, expert opinion, wine, product attributes.

JEL Classification: C23, C25, D12, H20.

---

*Bonnet: Toulouse School of Economics, INRA, University of Toulouse I Capitole; Hilger: U.S. National Marine Fisheries Service; Villas-Boas: Department of Agricultural and Resource Economics, University of California, Berkeley, CA, 94720. We thank Michelle Scizak, Reed Johnson, and Steve Flaming for help with the data; Will Liu and Jeff Coronado for help with the experimental implementation; and Kyle Birchard, Elizabeta Perova, Grant Chen, Patricia Javier, Katan Patel, and Elizabeth Creed for their excellent research assistance. The authors acknowledge the financial support received from the Giannini Foundation. We thank Peter Berck for comments. Corresponding author: Villas-Boas, sberto@berkeley.edu.

1 Introduction

There is a relatively large theoretical literature that discusses the ways in which informed consumers and expert opinion may solve adverse selection problems (Akerlof, 1970, Bagwell and Riordan 1991). Despite the growing body of empirical work that has applied this theory to the data (Hilger et al, 2011, Sorensen and Rasmussen, 2004, Reinstein and Snyder, 2005), there remains an empirical gap where actual consumer revealed choices are used to test the theoretical mechanisms that underlie the ways in which consumers make inferences about the quality of products with and without expert opinion information. This is particularly true for products where producers use observable attributes such as price and brand advertising to signal quality as theory predicts in an asymmetric information and repeated interaction context (see e.g., Bagwell and Riordan, 1991, Mahenc, 2004, Shapiro, 1983, Spence, 1976, and Tellis and Wernerfelt, 1987). The contributions of this paper are to first to investigate whether expert opinions, in the form of an expert score, affect demand for labeled and unlabeled wine products in the context of incomplete information about wine quality, and second to test for some economic mechanisms about the consumers belief and find evidence on how consumers use information about observable characteristics to infer the quality of products.

To analyze this question, we use purchase decisions from a retail field experiment in which we label a random subset of scored wine products in one treated store and do not label those same scored products in comparable control stores. Furthermore, we do not label in either the treatment or control stores the remaining scored products that were not selected, nor the products without available scores. We use a detailed monthly product level panel scanner data set for labeled and unlabeled products at treatment and control stores before and after the implementation of a shelf labeling field experiment. We combine the scanner data with detailed product specific attribute information readily available to all consumers through information on the product, such as wine brand, varietal, and region of production, as well as the price. Using revealed consumer preferences, the objective of this paper is first to estimate the average and heterogeneous effects of the field experiment on wine consumption of treated (scored and labeled) products, unlabeled products with scores, and unlabeled products without scores. Moreover, we investigate heterogeneous treatment effects according to product characteristics such as price range or regions of production. We also test for evidence of spillover effects among products of the same brands or same varietals to reveal possible mechanisms behind how consumers make inferences about the quality of products they buy given observable product attributes in the choice set.

This paper uses the same experimental design as Hilger et al. (2011), but extends the analysis to to data beyond treated (scored and labeled) wines to a larger set that includes unlabeled wines with scores, and unlabeled wines without scores. We match the treatment store to control stores, making sure all three subgroups of wines (treated, unlabeled with scores, and unlabeled wines without scores) share common pre period trends in the treatment and control stores. Our contribution is to extend Hilger et al. (2011) to an investigation of incomplete information through testing the possible mechanisms behind how consumers form beliefs about product quality when making their revealed choices given available information. While Hilger et al. (2011) documented a carefully identified causal positive average effect of expert opinions on wine demand, especially for higher scores, and no average effect on the unlabeled wines, we dig deeper in this paper to try and understand why this is so, and how consumers incorporate the new information in deciding what future choices to make. For instance, we test whether there are positive or negative spillover effects on sales of unlabeled products that belong to the same brand or varietal as labeled products.

Our estimates suggest there to be a positive and significant overall average effect and that demand increases more for higher score wines than for lower score wines (as in Hilger et al., 2011). New results are suggestive that higher scores matter more for prices in the lower quartile of the overall wine price distributions, which rejects consumers previously perceiving low price as signaling high quality. Our findings are instead consistent with consumer behavior where demand does not move for higher price quartiles once quality is revealed as consumers infer high quality for high prices. We also find positive spillover effects of this experimental treatment within brand for untreated wines, especially if the average score of wines within brand is high. However, if there is a higher percentage of wines within a given brand that are treated, that is the brand is intensively treated, we find that there are significant negative spillovers for untreated wines.

The remainder of the paper is structured as follows. Section 2 discusses the literature of consumer belief. Section 3 presents the experimental setting and the data. Section 4 the empirical strategy to estimate the effects of displaying labels on average and by product attributes to test theories of belief formations in the absence of quality information. Section 5 presents and discusses the findings and section 6 concludes.

2 Background Literature on Consumers Beliefs

There is a large empirical literature on understanding consumer beliefs about product quality through experimental analyses. In the drink and food markets, the perception of quality can be measured using two different scales: one for intrinsic attributes and one for extrinsic attributes (Jover et al., 2004). However, in the empirical modeling of real choice situations, only the extrinsic characteristics of the food products can be used to infer the quality of products. Among extrinsic characteristics, consumers may consider the price, brand name, country-of origin, and promotional activities of the products in their choice set (Ozretic-Dosen et al., 2007; Schnettler et al., 2009; Jover et al., 2004). Veale and Quester (2009) show that price and the country-of-origin effects could be more important than some intrinsic characteristics as taste in the wine market. Palma et al (2013) show that price is a key attribute in the contribution of expected quality by the consumer before tasting the product. Gustafson et al. (2016) uses an valuation experiment to show that consumers own knowledge affects preferences for wine products. However this literature is based on stated choices. The revealed choice approach has the advantage of 'face validity', as the data are consumers' actual choices when faced with real constraints on their own resources and the products available (Hensher et al., 1998; Whitehead et al., 2008). Consumers consider the internal costs and benefits of their potential choices and experience the consequences of their actions. Carson et al. (1996) shows through meta-analysis that estimates from stated and revealed preferences differ.

Empirical literature using revealed choices has analyzed the extent to which product quality information affects consumer behavior including branding (Montgomery and Wernerfelt, 1994), mandatory product labeling (Jin and Leslie, 2003; Kiesel and Villas-Boas, 2007), experimental labeling (Kiesel and Villas-Boas, 2007), and advertising (Ackerberg, 2003). Closely related to our paper, besides Hilger et al. (2011) are papers by Sorensen and Rasmussen (2004) for the book market, Reinstein and Snyder (2005) for the movie industry and Friberg and Gronqvist (2012) for the wine market. The key identification of the effects of expert opinion on movie demand in Reinstein and Snyder (2005) results from exploiting the timing of movie reviews by Siskel and Ebert. While they find no overall effect of reviews, they show that positive reviews increased box office revenues for narrowly-released movies and dramas, although it remains to be explained why. In the book industry, Sorensen and Rasmussen (2004) find that both positive and negative reviews in the New York Times increase book sales. Friberg and Gronqvist (2012) show that expert opinions affect the demand for wines in Sweden using five years of weekly data on sales, advertising, and expert reviews. A positive review raises the demand by more than 6% the week after but the effect gradually decreases and eventually becomes insignificant after twenty weeks. While all empirical evidence thus far seems to point out that expert opinions have a demand effect beyond just raising product awareness, the mechanism as to why remains to be explained.

Some theoretical works investigated those mechanisms. Price and brand advertising can be used by firms to signal quality in an asymmetric information and repeat interaction context (see Bagwell and Riordan, 1991, Mahenc, 2004, Shapiro, 1983, Spence, 1976, and Tellis and Wernerfelt, 1987). Shapiro (1983) and Bagwell and Riordan (1991) find that high prices signal quality if consumers attribute high quality to seeing a high price (as in Shapiro, 1983) and there is a proportion of informed consumers who buy if quality is good, unless there are too many uninformed buyers in the market.[1] If that is so, revealing high scores for high priced wines will have less effect than revealing high scores for lower priced similar wines. However, if a low price signals high quality through promotions to induce trial and repeated purchases once quality is assessed, for example as in Mahenc (2004), Spence (1976), and Tellis and Wernerfelt (1987), then high scores for high price wines may have a larger effect on demand than high scores for low price wines (that consumers had initially inferred as being good quality in prior to the additional label information). Gneezy et al. (2014) focused on price -dependant quality expectations with a theoretical model combined with lab experiments and show that expectations are important drivers of the price-quality relationship. Revealing information on a product could also lead to spillover effects on other products. Indeed, product brand information could improve the reputation of other products within the same brand. Costanigro et al. (2010) show that there exists a relationship between price and the structure of the name (collective and specific) and reputation in the wine market.

The objective of this paper is to empirically test theories of equilibrium price and quantity in the context of asymmetric information. For instance, we investigate whether the change in quantity purchased occurs among low or high priced wines. We apply those tests to the wine market where consumers choose the wine product that maximize their utility, where a

[1] In Shapiro (1983) to prevent adverse selection the incentive compatible equilibrium price-quality schedule (that prevents low quality to sell at high price and imitate being good quality) involves high quality products selling at a premium above their cost, so that this premium compensates the cost of providing the high quality product to all informed consumers and to a proportion of randomly purchasing uninformed consumers (Shapiro, 1983). Bagwell and Riordan (1991) show that the high quality producers price decreases as more consumers are informed.

product is a bundle of attributes. In the context of asymmetric information and a choice set that has wine products with different combinations of product attributes, consumers infer product quality based on the combination of those attributes and how they affect utility. In the asymmetric information literature, experts are considered informed consumers and the empirical test is whether uninformed consumers incorporate the information from experts when assessing the quality of goods. Wine is an experience good, in that consumers only realize it's quality after consumption, and sometimes even after consumption consumers may not really fully appreciate its quality (Ali and Nauges, 2007). There may a role for experts to reveal their opinion on wine product quality in order to help consumers make choices given the choice set of wine products with different prices, attributes, and the additional expert scores.

3 The Experimental Setting and the Data

Wine scores from a proprietary wine scoring system were displayed in the treatment store for four weeks during the month of April 2006. We use a proprietary score wine data set to randomly label 150 wine products with score labels that were attached below the price tag every week by the research team. It is important to note that not all wines with scores received a label. More specifically, 325 scored wines remained unlabeled. In addition, there remained nearly 613 wines in the choice set of consumers that were not scored by the experts. Therefore three subgroups of wines are present in the: 1) labeled wines with scores (14% of the choice set), 2) unlabeled wines with scores (28% of the choice set), and 3) unlabeled wines without scores (58% of the choice set).

Each label features the name of the proprietary scoring system and the wine's score that, in theory, ranges from the lowest 50 to the highest score of 100, but scores less than 70 are not released by the rating agency. Figure 1 displays the kernel density of the score histogram for labeled and unlabeled wines with scores in the treated store on the left panel and for the control stores on the right panel. Given a Kolmogorov-Smirnov (KS) test for equality of the distributions for both the labeled and unlabeled panels and a KS estimate of 0.0987 for the Treated Store with a p-value of 0.481 and a KS estimate of 0.0935 with a p-value of 0.518 for the Control Stores, we can not reject the equality of the score distribution between the labeled and unlabeled status of wines. There is therefore a nice match in the whole distribution of labeled and unlabeled wines in the treatment store and in the control stores which supports the difference in differences estimation strategy for the identification of the effect of score label.

The treatment store is located in a wealthy suburban neighborhood and is in the same marketing division as a set of 38 potential control stores.[2] The shared marketing division means that the pricing, promotion, and display layout is common among the treatment and control stores. This contributes to a good balance of observable determinants of quantities of wine sold that originate from the retail marketing strategy.

[2]Like in Hilger et al, (2011) on the one hand, if consumers in wealthy areas are likely to be more fully informed regarding wine quality than consumers in other areas, one could argue that we have a low likelihood of finding a significant treatment effect. On the other hand, if consumers in wealthier areas care more about quality and expert reviews, we would have a higher likelihood of finding an effect.

3.1 The Data

The data set consists of records for all wine products purchased in a certain week by store, from March 2003 until May 2006, which we will aggregate to the product by month by store level for our analysis. In particular, we observe purchase data by store and by week for 102 products among the 150 wine products for which we attached labels to the price tag every week. For the sub-sample of 325 scored wines that remained unlabeled, we observe purchase records for 230 products. In addition, we have purchase records for 610 of the 613 wines in the choice set of consumers that were not scored by the experts. Overall, we have purchase data for all wine products in the treatment store and a set of potential control stores, among which we will use only the four that best match the treatment store's pre period trends in labeled, unlabeled, and unscored wines, as we will show next.[3] The scanner data provide a unique wine product code identifier (UPC), the brand name and varietal of the wine, the number of bottles sold, the pre discount price paid, and any retail discount pricing offered. We aggregate the weekly sales data to the month-level for each unique product code and store to generate the number of bottles, average shelf price, average price paid (the shelf price net of discounts), and whether a bottle of wine was discounted. Pricing and discounting for each product are common for all the stores in the data and, moreover, wine pricing was not updated due to the selection of products into labeled and unlabeled status. Lastly, prices

[3] We use past data to match the treated store and possible controls based on similar pre trends not only for the labeled products but also for unlabeled scored produts and unlabeled unscored products unlike in Hilger at al (2011) where more control stores are used.

were not differentially updated in the treated store due to our experiment.[4]

For those wines for which proprietary wine score data exist, we merge the wine score data with the scanner data. Detailed product attribute data including: brand of the wine product, varietal (for example Cabernet or Merlot), aggregate wine type (red, white, or other), and origin of production and imported status is merged with the scanner data set.

We use data for multiple months before the treatment window to investigate pre period trends. As reported in Table 1, in the treatment store, we have data on 2,562 product month observations for treated wine products, 5,721 observations for untreated wine products with scores, and 16,194 observations for unscored wine products. In the four control stores the total number of observations for the three treatment status subgroups of wine products are 11,058, 24,578, and 72,871 respectively. The three product subgroups have the same proportion of observations across the treatment and control stores.

Summary statistics are reported in Table 1 and organized by store and wine treatment group as follows. In the first three columns we report descriptive statistics in the treatment store for treated/labeled wines in column 1, unlabeled wines with scores in column 2, and untreated wines without scores in column 3. Columns 4, 5, and 6 report descriptive statistics for the control stores for the respective three wine and score groups.

In the first row we report average quantity sold during the pre treatment month (March); this is followed by its standard deviation in the second row. Below that the average quantity sold during the treatment month (April), followed by its standard deviation. Average quan-

[4] While in our empirical setting the price does not change due to experimental labels, it is worth noting that an extensive stream of the literature is related to the effect of experts opinion on wine prices (Dubois and Nuages, 2010; Hadj Ali et al., 2010)

tity sold by product in pre treatment month is 17 bottles for labeled wines in the treatment store and 19.2 bottles in the control stores on average. Unlabeled wines with scores sell 25.6 bottles on average in the pre treatment month in the treated store and 29.1 bottles in the control stores on average. Finally unscored wines sell about 12 bottles by product in the treated store and 13.4 bottles by product in the control stores during the pre treatment month. Average prices in the pre treatment month, as well treatment month, and all treatment statuses are not statistically different from one another. Most of the wine consumers purchase is discounted in the both the pre treatment and treatment months: approximately 90 percent for treated wines (in both treatment and control stores), almost one hundred percent for untreated wines with scores, and approximately 75 percent for untreated wines without scores.

The next two rows of Table 1 reports the average and standard deviation of scores by treatment status for treated and control stores. Here we see that average scores at the treatment store are not statistically different between the treated wines and untreated wines, with scores of approximately 83 and standard deviations of roughly 3.2. This is corroborated by Figure 1's reporting of similar score distributions among treatment status for the treatment and the control stores.

The next two rows report the composition of red and white wines for each store and wine treatment group pair. For wines within the treatment store the treated group consists of 58 percent red wines, the untreated with score group consists of 45 percent, while the unscored group contains 50 percent red wine; a similar distribution characterizes the control stores. Similarly, white wine and other (not reported in Table 1) proportions by treatment status are similar between the treatment store and control stores.

3.2 Pre Period Trends in Treatment and Control Stores

To estimate the causal effect of revealing score information on a random subset of wines with scores on the three different subgroups of wine products as defined by treatment status, we need to verify that one crucial assumption is satisfied - that there are similar pre period trends for the treatment and the control stores used in the analysis.

First we investigate whether the pre period is balanced in terms of pre existing trends in total quantity sold of wine products for the treated store and control stores. Figure 2 presents the total quantity sold by store per month for the year preceding the treatment, where the treatment is indicated by vertical line. While the various stores differ in levels of quantity sold, their trends are quite similar. While the treatment store has a lower total quantity of wine sold than each control store, the trends for wine quantity purchases follow similar patterns in the treatment store and the four control stores.[5] Thus, while the treatment store has different quantities sold than the control stores, to the extent that these differences are constant over time, store fixed effects will control for all possible time invariant determinants of wine demand at the store level.

We repeat in Figure 2 the graphical analysis for the treated (labeled) products only, and, once again, while there is a lower quantity of wine products sold in the treatment store than in each controls store, the pre treatment trends follow each other closely. This pattern is also

[5]As a more rigorous test of parallel trends, we regress quantity on a time trend for the treatment and control stores separately. We find that the point estimates of the trend coefficients in treatment and control stores are not statistically different from each other. Furthermore, the time series correlation of the sample averages of treated and the control stores is high, suggesting that the treatment and control stores share broadly similar time varying patterns in the pre treatment period.

evident for untreated wines with scores and for the unscored wines, respectively. In sum, for all the subgroups, the trends are quite similar in the treated and control store which will allow us to investigate the causal effects of the treatment on treated wines, untreated wines with scores, and untreated wines without scores.

4 Reduced Form Empirical Specification

Using two data sets, a store level and a product characteristics data set, we estimate a reduced form revealed preference specification of consumer responses to the labeling treatment. In particular, we extend Hilger et al. (2011) measuring average and heterogeneous changes in wine consumption depending on wine characteristics to test theoretical predictions of consumer inferences about product quality, by taking advantage of the exogenous changes in information about product attributes introduced through the field experiment.

4.1 Global identification strategy

We follow a difference-in differences-approach (DID) commonly used in the policy evaluation literature (see Meyer 1995; Bertrand, Duflo, and Mullainathan 2004). Our control structure is twofold: temporal, as we compare one particular product's purchases in weeks with and without labels; and cross-sectional, as we compare purchases of similar types of products between the treatment and the selected control stores. We look at wine products in the three different subgroups defined by treatment status. That is, we look at the labeled products with scores, the unlabeled products with scores, and finally the unlabeled products with no scores.

Let a product i be defined as a wine option in a certain store. The DID model specification is as follows :

$$
\begin{aligned}
(1) \qquad Q_{igm} &= \alpha_i + \beta_1 \text{Price}_{igm} + \beta_2 \text{Treated Store}_{ig} + \beta_3 \text{Treatment Period}_m + \\
&+ \beta_4 (\text{Treated Store*Treated Period})_{igm} + \beta_5 \text{D}_{igm} + \epsilon_{igm}
\end{aligned}
$$

where Q_{igm} is the quantity sold of product i in product treatment status group g and month m, α_i is a product fixed effect, $Price_{igm}$ is the unit price of product i in product group g and month m, D_{igm} is the a discount dummy if the product i in product group g and month m is on sales, Treated Store_{ig} is an indicator for product i in treatment status group g being in the treated store. A time indicator Treatment Period_m is equal to zero in the pre treatment period and equal to one in the treatment period.

The coefficient of interest is of the interaction of Treated $Store_{ig}$ and the Treatment $Period_m$. When estimating equation (1) separately for g =(Treated Wines, Untreated Wines with Scores, Un-scored Wines), we obtain three coefficients of β_4, which are the average effect of the treatment on the wine demand of each of the subgroups of products defined by treatment status.

4.2 Heterogeneity and tests for Belief Updating

To investigate the heterogeneity of the effects according to observable product attributes we split data into four subsets in the treated status: high score and low price quartile, high score and high price quartile, low score and low price quartile, and finally low score and high price quartile. Score range allows testing if revealing higher scores are associated with larger demand responses. Price range allows testing for pre treatment consumer beliefs about wine quality. We then estimate the equation 1 for the four data subsets.

If the estimated β_4 for the high score low price subset is positive and significant, then we reject that consumers believed pre treatment that a low price signaled high quality (as in Mahenc, 2004, Spence, 1976, and Tellis and Wernerfelt, 1987). However, if β_4 is not significantly different from zero for the high score high price wines, we find evidence consistent with Shapiro (1983) and Bagwell and Riordan (1991) that similar to pre treatment beliefs, consumers believe that all else equal a high price means high quality. This also remains true if β_4 is not significantly different from zero for low scoring low price quartile wines. We do not have enough observations in the high price low score set to perform an empirical test for that group.

4.3 Test for spillover effects

When scores are revealed while holding constant other attributes, such as price (pre- and post treatment), brand, and varietal, consumers may update the way they use price in conjunction with other attributes as a quality signal in the post treatment period. We investigate whether there are some spillover effects due to belief updating, by observing consumers revealed preferences of unlabeled wines that share similar characteristics of labeled scored wines. In this last specification, we test for spillover effects by taking advantage of having collected varietal and brand name attribute data for all wines.

In particular, for each wine, we compute variable that measure the proportion of wines that are treated by brand and varietal. We also compute a variable defined as the average score of treated wines in the same brand of each wine, and a variable defined as the average score of treated wines in the same varietal. To test for spillovers within brands, we test whether or not the degree of treatment within a given brand allows increasing demand for wines in the same brand. If the coefficient is significantly different from zero we reject the null of no spillovers. We also perform the same investigation to test for spillovers within varietals. Lastly, we compare the estimates of brand and varietal spillovers.

5 Results

This section is organized as follows. First we present the average treatment effects on the treated wines. Then we turn to presenting heterogeneity estimates by score and by price range and testing belief updating about quality of wines. Finally, we look at spillovers among the treated wines and investigate whether there are spillover effects into unlabeled scored wines as well as unscored wines.

5.1 Average Treatment Effects on Treated Wines

We present the results from the reduced form specification of equation (1) in Table 2, where the dependent variable in the first two columns is the quantity of bottles sold of wine product i at store s and month-of-sample m and the ln of quantity in column (3). The first row contains the estimated treatment effects coefficient of the difference in differences interaction of the "Treated Store and Treated Period" dummy variables, rows two and three report the average quantity change for the treated store and the treated period, respectively. In column (2), we report on a specification with the addition of a price variable and a dummy variable that indicates whether or not the product is discounted, while Column (3) repeats the specification in column (2) with quantities and prices in ln's.

Treated products sell an average of 15 bottles a month and during the treatment period, and average demand drops for both the treatment and the control stores given the negative albeit not significant coefficient of -3.277 in levels in column (2) which corresponds to a 19.3 percent decline given in (3). As already show in Figure 2, average quantity is lower for the Treated store than for the controls, which is reiterated by the negative and significant point estimate of "Treated Store", by about 3 bottles, or 18.6 percent. The price coefficient is statistically different from zero and negative in all specifications. Given the ln-ln specification in column (3), we can interpret the coefficient on *Price* as a price elasticity. This coefficient equals -2.671, which ultimately means that demand for wine is elastic. Finally, average quantity sold increases by 55.7 percent when products are discounted.

In all three columns of Table 2, the coefficients on the interaction terms "Treated Store X Treated Period" are positive. For the specification in levels we see that revealing scores increased wine demand by 1.5 bottles, about 5 percent, and the increase is statistically different from zero, suggesting that wine product sales decreased less in the treated store in the treated period than in the control stores, leading to a positive average treatment effect. For the ln-ln specification the coefficient is 0.052, but not statistically different from zero.[6] Next, we show heterogeneity by score levels and along other observable attributes for different subsets of wines according to treatment status.

5.2 *Treatment effects by Score Levels on Treated Wines*

To further understand whether and how consumers change demand for different types of wine, we estimate equation (1) by interacting the treatment effect with the displayed level

[6]Our results are consistent with those of Hilger et al (2011) but the point estimates are not the same because we use a different set of control stores to estimate the average treatment effect here. In particular, we match the treatment store to the best controls in terms of pre period trends for not only the treated wines but also for the untreated wines with and without scores, and use four control stores.

of Score and present the estimates in column (1) of Table 3. In column (2) we estimate (1) for the subset of scores below the median score of 81 and in column (3) for wines with scores 81 and higher. We find that the higher the scores the larger the percent increase of bottles sold for treated wines. The coefficient of interest is the one for the row "Score Level X Treated Store X Treated Period" and we see that if the score displayed increases by one, then the number of bottles sold increases by 0.2 percent. Comparing columns (2) and (3) treatment effects, those in the row "Treated Store X Treated Period", we see that demand for wines below a score of 81 does not increase significantly (estimate of 0.005) while for scores 81 and higher demand increases significantly on average by 5.8 percent. In sum, we find a significant positive average consumer response to expert opinion labels for wine and significant demand increases for higher scoring wines.

5.3 Testing for Price and Wine Region Quality Belief in Pre-Period on Treated Wines

Tables 4 and 5 present estimates of equation (1) for treated wines with scores 81 or larger, and for treated wines with score below 81, respectively. Both tables allow differentiating per price range and wine regions: imported wines are in column (1), domestic wines are in column (2), where they are all California wines, the lowest priced quartile wines are in column (3) and the higher price quartiles in column (4). The coefficient of interest is, as before, the difference in differences interaction coefficient, the β_4 in equation (1) which corresponds to the row "Treated Store X Treated Period". The null is that demand for each of this subset of wines does not change significantly with our treatment.

For domestic wines in column (1), we only reject the null that there is no treatment effect for score 81 and larger given that the coefficient on the "Treated Store X Treated

Period" interaction, β_4, is positive and significant. Our treatment caused "Domestic" wines with scores 81 or more to increase by 14.7 percent, consistent with consumers updating their beliefs about domestically produced wines if they receive scores of 81 or higher. In contrast, the demand for imported wines in column (2) drops significantly by 26 percent due to the treatment for high scoring imported wines, suggesting that consumers switch away from high scoring foreign wines to domestic wines with high scores. This is consistent with consumers associating domestic wines with lower quality than imported wines – a belief that changes once the scores are revealed, leading demand for domestic wines to increase and demand for imported wines to drop.

From the estimates of columns (3) and (4) of Table 5 we see that the estimated β_4 for the high score low price subset is positive and significant, which leads us to reject that consumers believed pre treatment that a low price signaled high quality (as predicted by Mahenc, 2004, Spence, 1976, and Tellis and Wernerfelt, 1987). Indeed, treated high quality wines in the lowest price quartile increase by 18.1 percent due to the treatment. In addition, the estimated β_4 is not significantly different from zero for the high score high price wines in column (5). Taken together, we find evidence consistent with Shapiro (1983) and Bagwell and Riordan (1991), namely that consumers already believe in the pre treatment period that a high price means high quality. This is reinforced by the fact that the estimate of β_4 in column (3) of Table 5 is not significantly different from zero for low scoring low price quartile wines. Revealing low scores for imported and also for domestic wines does not cause a significant change in demand, given the non-significant estimates in columns (1) and (2) of Table 5.

5.4 Belief Updating due to Spill Over Effects within and across Treatment Status

We want to investigate here whether, when scores are revealed under the unchanged (pre and post experiment) price distribution of available choices, brand and varietal, consumers update the way they use price in conjunction with all other attributes in order to infer quality in the post period. We test for belief updating in the form of spillover effects on unlabeled wines.

We compute a variable that features the average score of treated wines in the same brand of each wine, and a variable that features the average score of treated wines in the same varietal, especially given that the wine bottles are organized on the shelves by varietals rather than by wine brand. The test for spillovers is to test whether the magnitude of treatment within brand causes demand to increase or not for wines in the same brand, and perform the same empirical analysis within varietal taking advantage of the variation in the intensity of treatment by brand and varietal and also of the variation in average scores for treated wines by brands and varietals.

There are a total of 375 wine brands, where 297 of them have none of their products treated. However, for the remaining 76 wine brands, we have 6.37 percent of treated wines within a brand on average, with a standard deviation of 15.6. There are 4 brands that have all their available wines treated and some that have no wine treated, and the remaining 72 brands have intensity of treatments ranging between 4.5% and 66%. The intensity of treatment also varies among the 14 different varietals: on average 6.99% of wines in a varietal are treated and four varietals have no wines treated while the remaining ten varietals intensity of treatment ranges between 4.76% and 22% of wines treated.

In terms of average scores, among the 10 varietals that get treated the average score ranges between a varietal with an average score of 81.7 and a varietal with an average score of 89 among its treated wines. In terms of brand level variation in average scores, we have one brand with an average score of its treated wines being a 78 and one other brand with the highest average score among its treated wines being a 89. In between, average brand level scores range from 79 to 88 for their treated products, and the brand level average score is 83.6 with a standard deviation of 2.65.

Table 6 investigates within brand spillovers and reports the estimates in four columns according to treatment status and score levels for each subgroup of wines in the sample. Column (1) has treated wines with scores lower than 81, column (2) has treated wines with scores 81 or higher, column (3) has untreated wines with scores 81 or higher and column (4) has the group of wines without any scores.

The coefficients of interest are the ones associated with "Treated Store-Period X Average Score by Brand" and also "Treated Store-Period X Percent Treated by Brand". Given that the coefficients are not significantly different from zero in column (1) for the average and intensity of treatment by brand rows, we reject there to be within brand spillovers for treated wines of low score level. However, for the high scoring treated wines, we find that there are positive spillovers. In particular, we find that a one percent increase in the intensity of treated wines in a brand portfolio significantly causes the number of wine bottles to increase given the significant point estimate of 96.9% as shown in column (2). Given that among the treated brands, the range of treatment by brand is 4.5% and 66%, if the percentage of wines treated by brand increased to 66%, wine bottles would increase by 63.9% for each wine in the same brand that got treated.

For the untreated group that has scores higher than 81, in column (3), we reject the null of no spillovers. The coefficient of average score of treated wines in the same brand is positive and significant meaning that the high quality spills over to untreated wines belonging to the same brand portfolio of high average scored brands. Moreover, the coefficient associated with the percent treated by brand is negative and significant. This implies that consumers perceive the wines that receive no labels as being of low quality. While brands that receive high average scores benefit from positive spillovers to its untreated brands with scores, the more wines that receive a score within a brand, then the more consumers perceive the wines receiving no labels as being of poor quality.

For the unscored wines in column (4) we reject the null of no spillovers. In fact, the estimates of average score and percent treated within brand spillover measures are both significant and positive. This is consistent with consumers inferring that if a product belongs to a brand that has in general high scores or is highly labeled then the unlabeled product must be good. As a result, they associate brand with high scores affiliation with high quality. Both columns (3) and (4) measure spillovers into sales of wines that do not receive a treatment label. While we see that having a product in its brand that received a higher average scores has positive and significant spillovers for both untreated products with scores (in column 3) and untreated products without scores (in column 4), it is quite interesting that we see differences between column (3) and column (4) in terms of spillovers by intensity of treatment. For the group that never received a score (group in column 4) we find that they benefit from high intensity of treatment within its brand portfolio, while in column (3) it is the exact opposite, given the negative and significant point estimate associated with the intensity of treatment within brand. These are different types of wines: the sample of wines in column (3) has scores higher than 81 and the sample in column (4) would likely get a score less than 81 if scored. As a take away it appears that the spillover effects from the brand level intensity of treatment benefit significantly the worse untreated wines relative to the higher scoring wines.

Table 7 presents the estimates used to test for spillovers within varietals. Table 7 is organized the same way as Table 6 and only the coefficients in column (2) reject there to be no spillovers. For the sample of wines with scores 81 or higher, we find that higher average scores within varietal, lowers demand for treated wines in the same varietal. Additionally, the higher the intensity of treatment within varietal, the larger the increase in demand for treated wines of scores 81 or higher.

6 Conclusion

We empirically estimate the effect on wine demand from revealing expert opinion information about the quality of a subset of products in the choice set of consumers. We not only estimate average effects, but we break up the effects by subsets of groups of labeled and unlabeled wines to investigate and test possible theoretical mechanisms as to how expert opinion labels can reveal the way consumers make inference about the quality and compare among products in the context of incomplete information.

First, we find there to be a positive and significant overall average effect of the changes in information provision in the form of expert opinion on consumer preferences for the wine products labeled relative to the control products. There is also heterogeneity in the estimated effects in that demand increases more for higher score wines than for lower score wines as in

Hilger et al (2011) using even more data on wine choices. We estimate there to be spillovers effects of this experimental treatment within brand for untreated wines. If wines do not receive a label they do significantly benefit from average scores within the same brand but, if many products are labeled for the same brand we find evidence consistent with consumers inferring that the unlabeled wines for the same brand are of low quality and demand drops. This is not the case for varietial.

This paper contributes to the empirical literature as it uses actual revealed preferences where price is perceived as being positively correlated with wine quality when consumers have incomplete information. We find that consumers update their beliefs of low prices signaling lower quality, all else equal, given that they increase demand for lower quartile priced wines that have a high score. This result is consistent with previous literature findings of wine tasting experiments and choice surveys (see e.g., Mastrobuoni, et al, 2014).

Our findings are also consistent with pre treatment consumers inferring high quality for high prices, given that once quality is revealed, demand does not move for those higher priced wine quartiles as in Bagwell and Riordan (1991) and Shapiro (1983). We also reject that low price signalled high quality in the pre-period because revealing scores increases demand for those wines. If consumers already perceived low priced wines as high quality (as in Mahenc, 2004, Spence, 1976, and Tellis and Wernerfelt, 1987), then revealing scores should not have significantly changed demand for those low price wines. Future work could consider the extent that firms would likely respond to the additional quality information when deciding prices and product assortments.

7 References

Ackerberg, D. 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination,' International Economic Review.

Akerlof, George A., "The Market for 'Lemons: Quality Uncertainty and the Market Mechanism," Quarterly Journal of Economics, LXXXIV (1970), 488500.

Ali, H. 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.

Bagwell, K., Riordan, M.H., 1991. "High and Declining Prices Signal Product Quality." American Economic Review 81: 224-239.

Baltas, G. 2001. Expert Opinion Labeling: Issues and Policies. European Journal of Marketing. 35(5/6):708-721.

Berning, J., Jill J. McCluskey, H. H. Chouinard, K. Manning, D. Sprott, and S.B. Villas-Boas, 2007. "Grocery Store Shelf Labels, Expert Opinion and Pricing Information, and Consumer Choice,"mimeo.

Bertrand M., E. Duflo, S. Mullainatha, 2004. "How much should we trust difference-in-differences estimates?", Quarterly Journal of Economics, 119(1), pp. 249-275.

Burton, S. and Andrews, J.C. 1996. Age, product expert opinion, and label format effects on consumer perceptions and product evaluations. The Journal of Consumer Affairs. 30(1):68-89.

Caswell, J.A., and D.I. Padberg. "Toward a More Comprehensive Theory of Food Labels.' American Journal of Agricultural Economics 74(1992):460-68.

Costanigro, M., McCluskey, J. J., and Goemans, C. (2010). The economics of nested names: name specificity, reputations, and price premia.American Journal of Agricultural Economics,92(5), 1339-1350

Dranove, D., Kessler, D., McClellan, M., and M. Satterthwaite. 2003. “Is More Information Better? The Effects of ‘Report Cards’ on Health Care Providers,’ Journal of Political Economy.

Drichoutis, A.C., Lazaridis, P., Nayga, R.M. Jr. 2005. Expert opinion knowledge and consumer use of expert opinion food labels. European Review of Agricultural Economics. 32(1):93-118.

Drichoutis, A.C., Lazaridis, P., Nayga, R.M. Jr. 2006. Consumers’ use of expert opinional labels: a review of research studies and issues. Academy of Marketing Science Review. 9:1-22.

Dubois, P.; and C. Nauges, 2010. Identifying the Effect of Unobserved Quality and Experts’ Reviews in the Pricing of Experience Goods: Empirical Application on Bordeaux Wine, International Journal of Industrial Organization, 28(3), 205212.

Frewer, L.J., C. Howard, D. Hedderley, and R. Shepherd. “Reactions to Information about Genetic Engineering: Impact of Source Characteristics, Perceived Personal Relevance and Persuasiveness.’ Public Understanding of Science 8(1999):35–50.

Friberg R., and E. Grnqvist, 2012. Do Expert Reviews Affect the Demand for Wine?, American Economic Journal: Applied Economics 2012, 4(1): 193211.

Gneezy, A., Gneezy, U. and Lauga, D.O., 2014. A reference-dependent model of the pricequality heuristic.Journal of Marketing Research,51(2), 153-164.

Golan, E., F. Kuchler, and L. Mitchell. Economics of Food Labeling. Washington, DC: U.S. Department of Agriculture, Agricultural Economic Report 793, December 2000.

Gustafson, C.R., Lybbert, T.J. and Sumner, D.A., 2016. Consumer knowledge affects valuation of product attributes: Experimental results for wine.Journal of Behavioral and Experimental Economics,65, 85-94.

Hadj Ali, H., S. Lecocq, and M. Visser, 2010. The Impact of Gurus: Parker Grades and en primeur Wine Prices, Journal of Wine Economics, 5(1), 22-39.

Hilger, J., G. Rafert, and S. Villas Boas, 2011. "Expert Opinion and the Demand for Experience Goods: An Experimental Approach," Review of Economics and Statistics, 93 (4): 1289-1296.

Ippolito, P.A., and A.D. Mathios. 1993. "Information, Advertising and Health Choices: A Study of the Cereal Market.' Rand Journal of Economics, 21(3):459-80.

Jin, G. Z. and P. Leslie. 'The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." Quarterly Journal of Economics 118(2003):409-451.

Jover, A. J. V., F. J. Llorns Montes,M. F. Fuentes, 2004. "Measuring perceptions of quality in food products: the case of red wine," Food Quality and Preference, 15: 453469.

Kiesel, K., and S. B. Villas Boas, 2007. Got Organic Milk? Consumer Valuations of Milk Labels after the Implementation of the USDA Organic Seal, Journal of Agricultural & Food Industrial Organization, 5(1).

Levy, Alan S., Sara B. Fein, and Raymond E. Schucker. 1991. "Expert opinion Labeling Formats: Performance and Preference.' Food Technology, 45 (July), 116-121.

Mahenc, P. 2004. "Influence of informed buyers in markets susceptible to the lemons problem." American Journal of Agricultural Economics, 86(3):649659.

Mastrobuoni, G., F. Peracchi, and A. Tetenov, 2014. "Price as a signal of product quality: Some experimental evidence," Journal of Wine Economics, 9: 135-152.

Meyer, B. D., 1995. Natural and Quasi-Experiments in Economics, Journal of Business & Economic Statistics, 13(2), 151-161

Montgomery, C.A., and B. Wernerfelt, 1992. Risk reduction and umbrella branding, Journal of Business, 1, 31-50.

Ozretic-Dosen, D., V. Skare, Z. Krupka, 2007. "Assessments of country of origin and brand cues in evaluating a Croatian, western and eastern European food product," Journal of Business Research, 60: 130-136.

Palma, D., J. D. Ortuzar, G. Casaubon, L.I. Rizzi, and E. Agosin, 2013. Measuring Consumer Preferences Using Hybrid Discrete Choice Models. AAWE WORKING PAPER No. 137.

Reinstein, D. A. and Snyder, C. M. (2005), The influence of expert reviews on consumer demand for experience goods: a case study of movie critics. The Journal of Industrial Economics, 53: 2751.

Schnettler, B., R. Vidal, R. Silva, L. Vallejos, N. Seplveda, 2009. "Consumer willingness to pay for beef meat in a developing country: The effect of information regarding country of origin, price and animal handling prior to slaughter." Food Quality and Preference, 20: 156-165.

Shapiro, C., 1983. "Premiums for high quality products as returns to reputations." The Quarterly Journal of Economics, 98(4):65979.

Sorensen, A. T., and S.J. Rasmussen, 2004. Is Any Publicity Good Publicity? A Note on the Impact of Book Reviews, mimeo .

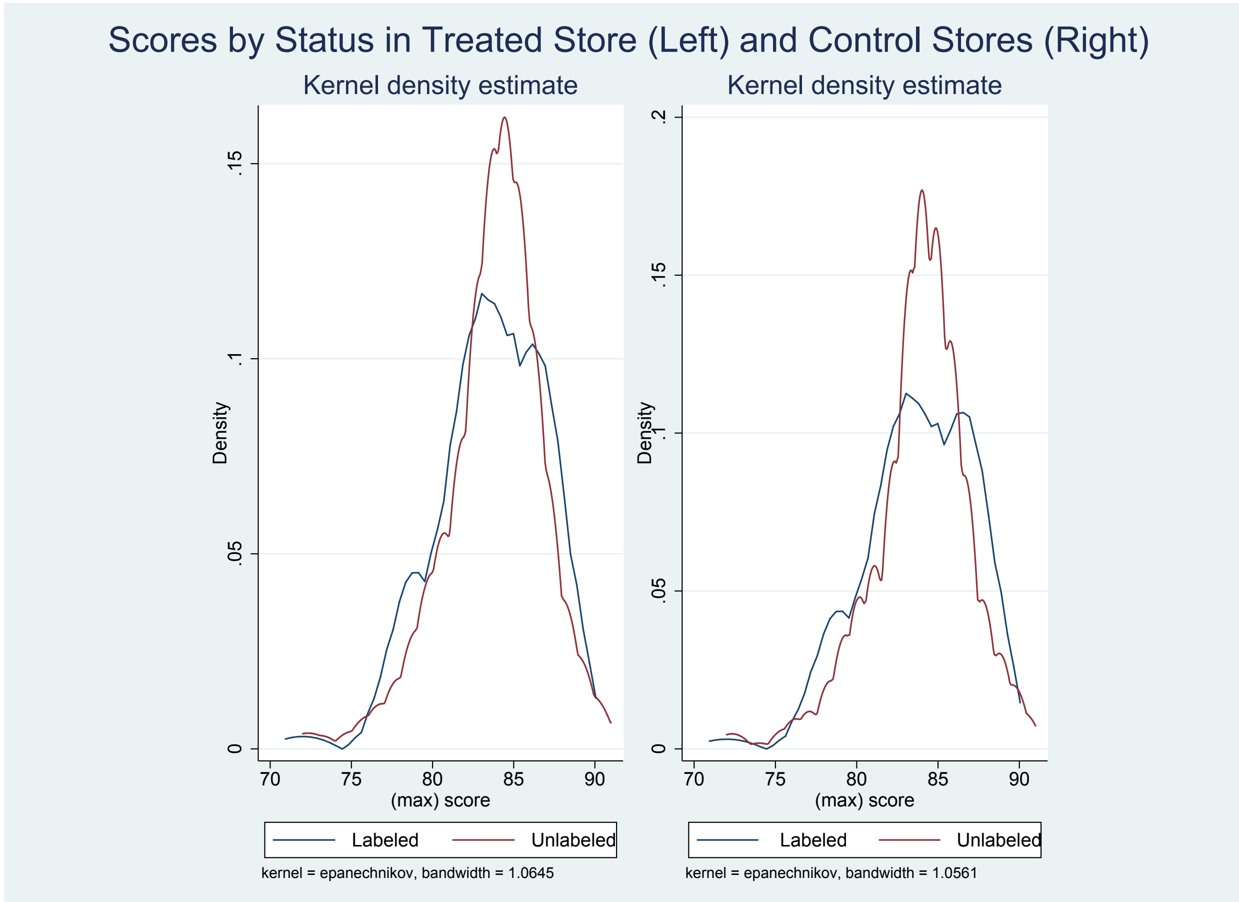
Spence, M. 1976. "Informational aspects of market structure: An introduction." The Quarterly Journal of Economics, 90(4): 591597.

Tellis, G. J., and Wernerfelt, B., 1987. “Competitive price and quality under asymmetric information.” Marketing Science, 6(3): 240253.

Tirole, J. 1988. The theory of Industrial Orgnaization, MIT press.

Veale, R., P. Quester, 2009. “Do consumer expectations match experience? Predicting the influence of price and country of origin on perceptions of product quality.” International Business Review, 18: 134-144.

Figure 1: Histogram of Scores by Status of Wines in Treated and Control Stores



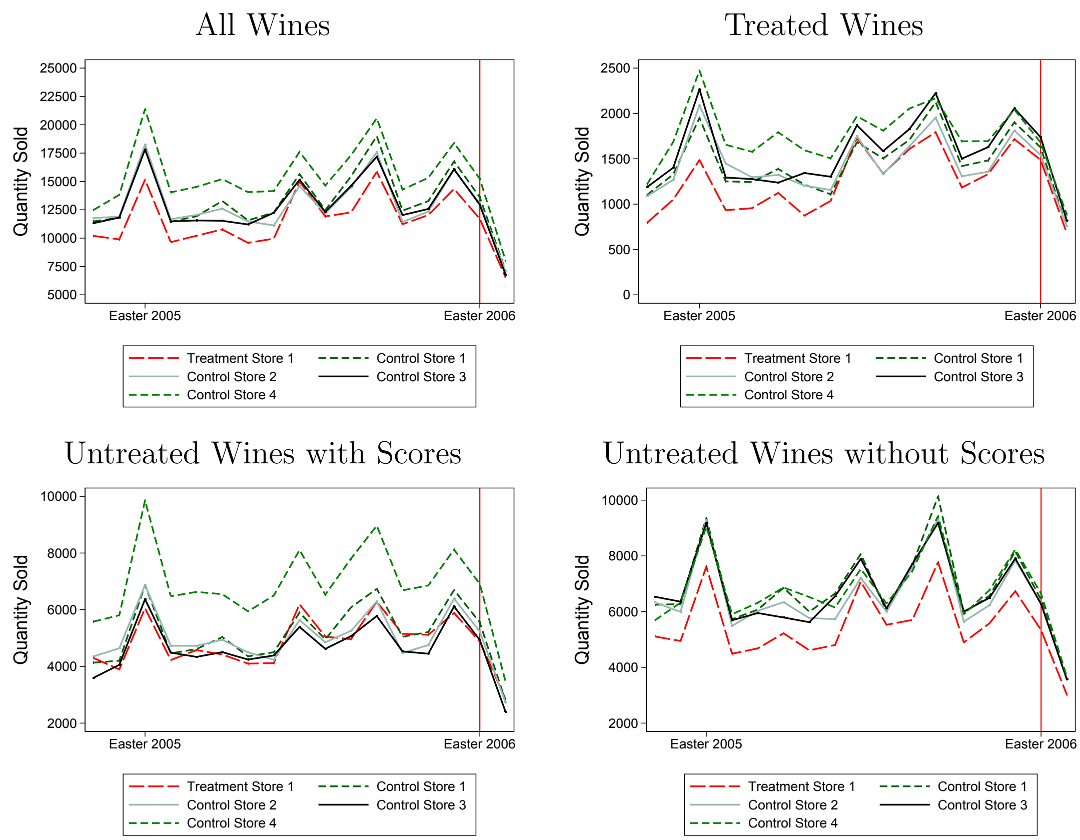
Note: There are scores available for wines pertaining to two status: Labeled status means these wine products are labeled and treated in the treated store. Unlabeled Status are the wines that have scores but do not receive a label with a score, thus are not treated in the Treated Store. The left panel displays the kernel density estimates of the score distribution for labeled and unlabeled wine products int e treated store. The right panel displays the kernel density estimates of the score distribution in control stores for the same group of wine products labeled and unlabeled (receiving a label and unlabeled in the treated store), given that we can see the same products in the control stores, which will be our counterfactuals in the difference in difference estimation strategy. Kolmogorov-Smirnov (KS) test for equality of the Distrib in each of the panels cannot reject the equality of scores distribution in the treated and also in the control stores between the labeled and unlabeled status of wines , given that KS (pvalue) is for the Treated Store= 0.0987 (0.481) and for the Control Stores =0.0935 (0.518).

Table 1: Summary Statistics of Wines by Treatment Status for Treated and Control Stores

	(1)	(2)	(3)	(4)	(5)	(6)
	Treated Store			Control Stores		
	Treated Wines	Untreated Wines With Scores	Untreated Wines Without Scores	Treated Wines	Untreated Wines With Scores	Untreated Wines Without Scores
Quantity (March)	16.99	25.60	11.97	19.19	29.06	13.42
	(26.01)	(44.90)	(21.57)	(26.01)	(53.28)	(23.75)
Quantity (April)	14.88	22.08	10.31	16.93	24.56	11.21
	(22.23)	(40.35)	(17.96)	(22.53)	(47.45)	(19.70)
Price (March)	10.98	10.15	10.98	11.23	10.38	11.36
	(5.00)	(4.81)	(6.43)	(5.02)	(4.91)	(6.50)
Price (April)	10.96	10.07	10.93	10.97	10.31	11.38
	(5.15)	(4.73)	(6.14)	(4.83)	(5.13)	(6.40)
% discounted (March)	0.91	0.99	0.75	0.88	0.99	0.74
% discounted (April)	0.88	0.99	0.78	0.89	0.98	0.75
Score	83.13	83.65		83.05	83.68	
	(3.20)	(3.20)		(3.29)	(3.01)	
% red	0.58	0.45	0.50	0.59	0.46	0.51
% white	0.35	0.48	0.40	0.34	0.46	0.39
Average Score by Brand						
Average	83.05					
Min - Max Score	78 - 89					
% Treated by Brand	6.37					
Min - Max % Treated	0-100					
Average Score by Varietal						
Average	83.59					
Min - Max	81.7 - 89					
% Treated by Varietal	6.99					
Min - Max % Treated	0 - 22					
Number Brands	375					
Number Varietals	14					
Number Wines	101	230	563	102	235	599
Number Observations	2562	5721	16194	11058	24578	72871

Standard Deviations in parentheses. First three columns for Treated Store, next three for Control stores.

Figure 2: Monthly Quantities Sold by Store



Note: Treatment period is the Vertical Red Line.

Table 2: Difference-in-Difference: Average Effect of Expert Opinion Labels on Retail Wine Sales

	(1) Quantity Sold (Q)	(2) Q	(3) ln(Q)
Treated Store X Treated Period	1.520***	1.507***	0.052
	(0.245)	(0.240)	(0.046)
Treated Store	-3.104***	-3.419***	-0.186***
	(0.261)	(0.266)	(0.015)
Treated Period	-1.851	-3.277	-0.193
	(3.345)	(2.797)	(0.171)
Price		-2.756***	
		(0.127)	
Discount Dummy		2.717***	0.557***
		(0.592)	(0.048)
ln(Price)			-2.671***
			(0.135)
Mean of Dep. Variable	15.118	15.118	1.943
Num of Obs.	13619	13619	13617
R squared	0.462	0.498	0.595
Product FE	X	X	X

Clustered errors in parentheses at the month level. Controls are best 4 stores. The dependent variable is the quantity by product sold per month except the last column, which is ln quantity sold.
$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 3: Difference-in-Difference: Score Level Effect of Expert Opinion Labels on Retail Wine Sales

	(1)	(2)	(3)
	ln(Q)	ln(Q) and Score < 81	ln(Q) and Score > 80
ln(Price)	-2.671***	-2.976***	-2.590***
	(0.135)	(0.217)	(0.150)
Discount Dummy	0.556***	0.481***	0.572***
	(0.048)	(0.051)	(0.055)
Score Level X Treated Store X Treated Period	0.002***		
	(0.000)		
Treated Store X Treated Period	-0.060	0.005	0.058**
	(0.060)	(0.156)	(0.025)
Treated Store	-0.186***	-0.248***	-0.172***
	(0.015)	(0.029)	(0.018)
Treated Period	-0.193	-0.317*	-0.165
	(0.171)	(0.163)	(0.171)
Mean of Dep. Variable	1.943	2.154	1.893
Num of Obs.	13617	2609	11008
R squared	0.595	0.553	0.600
Product FE	X	X	X

Clustered errors in parentheses. Clusters are at the month level. The dependent variable is ln quantity in all the other columns.

In Column 3 only wines with Scores Less than 81, and Column 4 Wines with Scores greater and Equal to 81.

$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 4: Heterogeneous Score Level Effect of Expert Opinion Labels on Retail Wine Sales Scores 81 or Higher

	(1) Imported	(2) California=Domestic	(3) First Quartile Price	(4) Highest Quartiles Price
ln(Price)	-2.884***	-2.539***	-2.873***	-2.554***
	(0.194)	(0.167)	(0.403)	(0.137)
Discount Dummy	0.425***	0.605***	0.667***	0.561***
	(0.070)	(0.062)	(0.094)	(0.056)
Treated Store X Treated Period	-0.261***	0.147***	0.181***	0.039
	(0.025)	(0.034)	(0.047)	(0.025)
Treated Store	-0.123***	-0.183***	-0.231***	-0.162***
	(0.031)	(0.022)	(0.045)	(0.017)
Treated Period	-0.177	-0.161	-0.231	-0.156
	(0.152)	(0.175)	(0.157)	(0.173)
Mean of Dep. Variable	1.699	1.940	2.387	1.813
Num of Obs.	2123	8885	1545	9463
R squared	0.612	0.595	0.567	0.595
Product FE	X	X	X	X

Clustered errors in parentheses. Clusters are at the month level. The dependent variable is the ln quantity in all columns and only featuring Wines Scored 81 of More. In Column 1 imported wines, in column 2 domestic (all California) Wines, Column 3 Lowest Quartile Priced Wines, Column 4 higher Quartile Price Wines.
$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 5: Heterogeneous Score Level Effect of Expert Opinion Labels on Retail Wine Sales Scores below 81

	(1) Domestic	(2) Imported	(3) First Quartile Price	(4) Higher Quartiles Price
ln(Price)	-1.423**	-3.396***	-3.287***	-2.571***
	(0.646)	(0.184)	(0.208)	(0.404)
Discount Dummy	0.750**	0.403***	0.449***	0.545***
	(0.309)	(0.043)	(0.092)	(0.081)
Treated Store X Treated Period	1.068***	0.586	-0.050	0.093
	(0.137)	(0.620)	(0.253)	(0.088)
Treated Store	-0.363**	-0.158***	-0.287***	-0.206***
	(0.159)	(0.008)	(0.050)	(0.057)
Treated Period	-1.034***	-0.152	-0.563***	-0.086
	(0.263)	(0.161)	(0.148)	(0.191)
Score Level X Treated Store X Treated Period		-0.008		
		(0.007)		
Mean of Dep. Variable	1.690	2.021	2.579	1.765
Num of Obs.	151	6489	1245	1364
R squared	0.218	0.579	0.553	0.427
Product FE	X	X	X	X

Clustered errors in parentheses. Clusters are at the month level. The dependent variable is the ln quantity in all columns and only featuring Wines Scored Less than 81. In Column 1 imported wines, in column 2 domestic (all California) wines in column 3 Lowest Quartile Priced Wines, Column 4 higher Quartile Price Wines. Not enough observations in this sample for Imported wines.
$*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 6: Spillover Effects of Expert Opinion Labels on Retail Wine Sales Within Brand

	(1)	(2)	(3)	(4)
ln(Price)	-2.655***	-2.981***	-0.916***	-2.055***
	(0.166)	(0.220)	(0.265)	(0.083)
Discount Dummy	0.587***	0.481***	0.436***	0.446***
	(0.060)	(0.052)	(0.068)	(0.039)
Treated Store X Treated Period	0.172	-2.593	0.191	-0.291***
	(0.547)	(1.832)	(0.350)	(0.081)
Treated Store-Period X Avg Score by Brand	-0.001	0.031	0.022***	0.003**
	(0.007)	(0.021)	(0.006)	(0.001)
Treated Store-Period X Percent Treated by Brand	-4.764	96.941**	-2198.940**	25.525***
	(4.592)	(39.626)	(962.632)	(5.334)
Treated Store X Average Score by Brand	-0.002	0.012**	0.002	0.000
	(0.002)	(0.005)	(0.003)	(0.001)
Treated Store X Percent Treated by Brand	6.377*	-18.342	-28.109**	-8.399
	(3.239)	(31.236)	(12.723)	(5.415)
Treated Period X Percent Treated by Brand	4.878	-52.793**	-48.759*	4.279
	(5.935)	(20.305)	(22.671)	(14.757)
Treated Period X Average Score by Brand	-0.005	-0.008**	0.001	-0.002***
	(0.004)	(0.003)	(0.004)	(0.001)
Treated Store	0.007	-1.157***	0.107	-0.162***
	(0.195)	(0.340)	(0.156)	(0.051)
Treated Period	0.226	0.390	-0.276	-0.014
	(0.469)	(0.420)	(0.261)	(0.172)
Mean of Dep. Variable	1.953	2.154	1.040	1.716
Num of Obs.	9550	2609	701	24270
R squared	0.606	0.557	0.278	0.604
Product FE	X	X	X	X

Clustered errors in parentheses. Clusters are at the month level. The dependent variable is the ln quantity in all columns. Column (1) has treated wines with scores lower than 81, Column (2) has treated wines with scores 81 and higher, Column (3) has untreated wines with scores 81 and higher, Column (4) has untreated wines without scores (for which we do not have available score information from experts. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 7: Spillover Effects of Expert Opinion Labels on Retail Wine Sales Within Varietal

	(1)	(2)	(3)	(4)
ln(Price)	-2.600***	-2.973***	-1.266***	-2.082***
	(0.154)	(0.220)	(0.162)	(0.057)
Discount Dummy	0.572***	0.482***	0.351***	0.423***
	(0.055)	(0.052)	(0.044)	(0.027)
Treated Store X Treated Period	0.491	2.422***	4.671	0.110
	(0.492)	(0.480)	(3.074)	(0.145)
Treated Store-Period X Average Score by Varietal	-0.004	-0.042***	-0.072	-0.001
	(0.008)	(0.012)	(0.057)	(0.003)
Treated Store-Period X Percent Treated by Varietal	-96.594	811.139*	-171.502	-44.117
	(70.401)	(376.050)	(667.436)	(79.485)
Treated Store X Avg Score by Varietal	-0.003**	-0.010*	-0.011	-0.002***
	(0.001)	(0.005)	(0.011)	(0.001)
Treated Store X Percent Treated by Varietal	42.731**	145.838	-459.277*	91.952***
	(17.085)	(168.747)	(221.059)	(9.672)
Treated Period X Percent Treated by Varietal	148.570***	-207.558	-251.699	38.832**
	(31.031)	(122.851)	(265.468)	(13.484)
Treated Period X Avg Score by Varietal	-0.004	0.006	0.047	-0.005***
	(0.003)	(0.009)	(0.029)	(0.001)
Treated Store	0.027	0.350	1.421**	-0.098**
	(0.098)	(0.325)	(0.625)	(0.035)
Treated Period	-0.051	-0.538	-3.339*	0.140
	(0.109)	(0.434)	(1.829)	(0.139)
Mean of Dep. Variable	1.893	2.154	0.983	1.592
Num of Obs.	11008	2609	943	90588
R squared	0.601	0.554	0.278	0.559
Product FE	X	X	X	X

Clustered errors in parentheses. Clusters are at the month level. The dependent variable is the ln quantity in all columns. Column (1) has treated wines with scores lower than 81, Column (2) has treated wines with scores 81 and higher, Column (3) has untreated wines with scores 81 and higher, Column (4) has untreated wines without scores (for which we do not have available score information from experts. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$