Random Utility Models, Wine and Experts

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

July 12, 2020

Abstract

We empirically investigate the welfare effects of providing product quality information to consumers in the form of expert opinion scores in a setting of asymmetric information. Identification of the effects of the provision of information comes from a field experiment in the retail wine market. We use a monthly-product-store panel scanner data set, collected before and during the field experiment, which involves treating a random subset of wine products by displaying expert scores in one store, and comparing sales in similar non-treated stores. Using a structural random utility model of demand, we show that, on average, consumers significantly value one score point increase due to the treatment by about 0.5 to 0.7 cents. As a consequence, for a bottle featuring an average score of 83, consumers would be willing to pay additionally between 20 and 60 cents more due to the treatment. Using counterfactual scenarios, we find that adding expert opinion shelf labels increases consumer surplus. In addition, allowing for strategic price reactions by retailers would lead to an overall significant welfare improvement given that (i) consumers significantly value the score information albeit facing higher prices, and (ii) the profits increase with the market power of firms.

Keywords: Random utility models, Structural random coefficients discrete choice models, Field experiment, Labels, Information, Expert opinion, Wine, Product attributes, Willingness-to-pay, Welfare Changes, Counterfactual simulations, Structural supply models, Product differentiation, Market power.
JEL Classification: M30, Q18, Q25, Q21, C12, C24.

This paper uses a randomly assigned introduction of expert opinion scores into the wine market to estimate demand and infer the implied revealed preference willingness to pay (WTP) for expert opinion information in the form of scores and implied marginal costs of suppliers under different wine pricing scenarios. Using the structural demand and supply model estimates we simulate welfare changes due to adding expert opinion labels into the market.

We collect the data by designing and implementing a choice experiment in one retail store, where we display expert opinion score labels for a random subset of wines across four weeks. We take advantage of unique data set comprised of retail scanner data and a field experimental data set, and build on the methodological breakthroughs that have arisen in the discrete choice literature when analyzing consumer demand (see McFadden 1999 and Train 2002 for a survey). Using a field experiment, we combine evidence from the revealed preference variation in observed choices to specify and estimate flexible random coefficient logit demand models and supply pricing models.

Our research goal is to empirically investigate the welfare effects of revealing product quality to consumers in the form of expert opinion score information, in a setting of asymmetric information such as the wine market, where consumers know less than producers about the quality distribution of products in their available choice sets. In order to investigate the welfare effects due to expert opinion shelf labels being implemented, we specify a demand and supply model and consider extreme cases for the price decisions of suppliers. First, we assume that there is perfect competition, with prices set to marginal costs, allowing the estimation of changes in welfare with and without shelf labels. In this case, prices are not adjusted in the counterfactual expert score scenario. Second, we consider the other

extreme case where suppliers have market power. Given demand and supply side behavior cases, we can bound marginal costs of the wine products. Assuming that prices are chosen in a Bertrand Nash fashion, we recover an upper bound on marginal costs by subtracting from prices the estimated margins under firms' Bertrand Nash pricing. Assuming monopoly pricing, we recover a lower bound on products' marginal costs as the difference between observed prices and implied monopoly markups. Using marginal cost estimates and demand preferences and allowing for strategic reactions by retailers, we estimate welfare gains for revealing scores from counterfactual scenarios.

Given asymmetric information on product quality, consumers must infer quality based on observable attributes at the time of purchase. High quality is typically positively correlated with higher average prices in many markets (Rao 2005; Shiv, Carmon, and Ariely 2005). Evidence from blind tastings in the wine market indicates that consumers attribute a positive premium to wines that are perceived as higher quality. Bonnet, Hilger, and Villas-Boas (2020) shows that uninformed consumers' purchases are consistent with beliefs that high quality is positively related to wine prices. In many markets, experts provide additional insight about the quality of products they evaluate, and develop expert ratings or scores that are commonly available to consumers. Producers value expert scores and opinions if they are able to charge higher prices for their high quality products, as they use the higher scores as a product differentiation device to increase market power (e.g., for wine see Ali, Lecocq, and Visser 2008). On the contrary, whether consumers value expert opinion information in this setting remains an unanswered question. Following the advertisement literature as conceptualized in Johnson and Myatt (2006), specifying how the shelf labels enter the random utility demand model, we consider that the expert scores shelf labels can

have either a role of creating a "hype," by influencing directly the indirect utility in the form of a demand shifter for the products receiving the label in a similar fashion, or can have an informative role, by influencing how consumers value an advertised attribute among the products in the choice set.

The first contribution of this paper is to provide structural demand estimates of wine products as a function of wine attributes, such as price, region of production, varietal, color, brand, and expert opinion scores, using revealed choices through scanner purchase data. 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, Louviere, and Swait 1998; Whitehead et al. 2008). Consumers consider the internal costs and benefits of their potential choices and experience the consequences of their actions. Choices based on perceived costs and benefits better reflect the values of the population and allow for more valid estimates of willingness to pay. Carson et al. (1996) shows through meta-analysis that estimates from stated and revealed preferences differ. Previous work using the same wine experiment data estimated the reduced form demand effects of wine scores (Hilger, Rafert, and Villas-Boas 2011). Subsequent work by Bonnet, Hilger, and Villas-Boas (2020) sought to identify the mechanisms underlying these demand effects, attributing value to different quality grades through the use of a 75 to 100 numeric score. The lack of previous research estimating the value of expert scores, in wine or other markets, is related to the challenge of identifying unbiased demand responses to scores that are uncorrelated with other strategic decisions taken by firms, such as pricing, branding, and product portfolio assortment choices. This challenge is circumvented in this paper, given that the treatment of wines (through revealing their expert scores) is randomly assigned across

the potentially scored wine products and is uncorrelated with marketing variables of the wine producers and the retailer. The treated wines had their scores displayed though a label placed on the supermarket shelf underneath the product price tag. When we ran the experiment, we ensured that neither the retailer, nor the producing vineyards, adjusted their marketing variables to take into account the experimentally disclosed score information at the point of purchase.

The second contribution of the paper is to estimate welfare effects, in terms of changes in consumer willingness to pay (WTP) and changes in firm profits, due to expert opinion score treatment. We measure welfare changes due to the introduction of expert opinion information, as a reduction of asymmetric information in the wine market. We develop and estimate a structural model (as in Berry, Levinsohn, and Pakes 1995; Nevo 2000) of wine demand, specifying a flexible random utility choice framework (as in McFadden 1974; McFadden 1981; Train 2002). That is, we analyze actual response behavior within a designed field experiment for wine retail products to directly estimate the revealed preferences and corresponding welfare effects of expert opinion scores as additional wine product attributes. In so doing, we will provide industry participants and policy makers with important information on the efficacy of score based quality-labels, as well as information on consumers' actual wine preferences and consequences for firms' ability to change prices given displayed expert opinion labels (Ali, Lecocq, and Visser 2008).

Related empirical literature 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), consumer peer-user ratings (Anderson and Magruder

2012; Jacobsen 2015), and advertising (Ackerberg 2001; Ackerberg 2003). Closely related to our paper, besides Hilger, Rafter, and Villas-Boas (2011), which estimates reduced form effects of the same experiment of displaying expert scores of wine, are papers by Berger, Sorensen, and Rasmussen (2010) on the book market and Reinstein and Snyder (2005) on the movie industry. The key identification of the effects of expert opinions 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, Berger, Sorensen, and Rasmussen (2010) find that both positive and negative reviews in the New York Times increase book sales. Our major contribution, extending all previous works, is that we are the first to assess demand side valuation of expert opinion labels using actual point of purchase decisions of consumers in a field experiment setting. Indeed, we use wine purchases in a store of a California retail chain before and after the label introduction as well as wine purchases in other stores of the same retail chain as a control for other demand shocks. We utilize a flexible discrete choice model (e.g. Berry, Levinsohn and Pakes 1995; McFadden and Train 2000; Nevo 2000; Nevo 2003; Swait and Andrews 2003) that incorporates heterogeneity in demand. The framework allows the empirical testing of the null hypothesis that the displayed expert opinion scores in the treatment labels are not valued by consumers. We estimate the value of the scores due to our intervention at the treated store, and as a result what we estimate does not reflect the full value of the score of wine quality.

Our demand estimates suggest that consumers would need to be offered a discount if wines feature a score less and equal than 71 and that, on average, they value a score point

increase by about 0.5 to 0.7 cents due to the treatment. As a consequence, for a bottle featuring an average score of 83 consumers would be willing to pay additionally between 20 and 60 cents more if the score was revealed to them in the form of a label. In terms of wine regions, consumers are on average willing to pay 2 dollars more for California wines relative to all other wine regions together and there is substantial variation in estimated preferences for wine varietals. We use the demand estimates and estimated supply margins and marginal costs under different supply scenarios to investigate what are the welfare changes due to adding our expert opinion score experimental labels. The baseline prices are prices without scores being revealed, which happens in the pre-period in the treated store. Then we simulate the new equilibrium prices that result from adding labels under Bertrand Nash and under monopoly pricing in the treated store in the pre-period. We estimate changes in choices that imply significant consumer surplus gains, which is due to the fact that consumers value scores positively and also given that firms setting prices as Bertrand Nash do not significantly change their prices given revealed expert opinion scores in the shelf labels. However, in the monopolist case, when revealing expert opinion scores that are positively valued by consumers, firms are able to raise prices significantly and welfare net effects are lower than in more competitive supply scenarios. We then find a significant net welfare gain of adding expert opinion labels under competition (as consumers value the information and switch to wines receiving scores) and a significant net welfare gain under Bertrand Nash (given that prices do not change much in this case), and a lower welfare gain under the least competitive supply setting of a monopolist seller (given that prices increase significantly on average). Overall we estimate that adding expert score labels leads to a significant welfare gain representing about eight percent of total revenue from wine sales in this store.

The remainder of the paper is structured as follows: In the next section we present the data and experimental variation used. Next we specify the structural demand model, derive how to obtain the implied estimates of consumer valuation for expert scores using macro-level aggregate choice data for treated and control stores and discusses the structural estimates. In the following section, we specify and estimate consumer surplus changes, profit changes and total welfare changes under alternative supply scenarios and policy scenarios of changing the expert scores shelf labels that consumer see when making choices. Lastly, we discuss implications of our findings and provide closing remarks.

The Experiment and the Data

In April 2006, wine ratings from a proprietary wine scoring system were displayed in one retail grocery store for four weeks, henceforth defined as the treatment store. We randomly selected 127 products to be labeled among wine products with scores, which corresponds to displaying scores for about 20% of the wines in the consumers' choice set. Each label featured the name of the proprietary scoring system and the wine's score. In theory, wine scores range from 0 up to the highest score of 100; however, scores less than 70 are not released by the rating agency.

The treated store is in the same marketing division as a set of potential control stores. The pricing, promotions, and display layouts are common among all of the stores in a marketing division, leading to a good balance of observable determinants of quantities of wine sold, originating from the retail marketing strategy.

The choice of the treated store was made at random among a subset of stores of the same

retail chain, within driving distance from the residence of the researchers, as the treatment had to be done manually by us. During the first week, we tag all the treated wines in the treated store in the evening of a Tuesday. The second week, also on the evening on the Tuesday, we visited the treated store again to make sure all our labels were still up and correctly showing. They always were. The only time we had to reattach a new label tag was when a certain product had a price change (which was set in advance and independently on our treatment). In the control stores, the same products also had the same price change, based on the usual price schedule set in advance and common to the treatment and control stores. We want to be clear that the prices did not change because we labeled, the prices change according to the pricing and discounting strategy set by the central division. The treated retail store is in northern California, and we use the set of stores that exhibit pre period trends similar to the treated store. Controls are also located in the same North California region, but not closely located to the treated store to avoid consumers possibly shopping at both treatment and control stores. The type of store is a typical retail store, it is not a wine specialty store, and it is also not a hard discount retailer. The treated and control stores in the data have similar store square footage, and the same store format, marketing promotions, and displays. Finally, control stores used in the empirical analysis also share similar socio-economic characteristics to the treated store in terms of zip code level census observables.

Figure 1 displays the kernel density of the score distribution for treated wines in the treated store and the kernel density of the score distribution of the unlabeled products sold in the control stores, given that we can see the same products in the control stores. Given a Kolmogorov-Smirnov (KS) test for equality of the distributions, we cannot reject that the

distribution of scores are equal across the treated and control stores. Therefore, there exists a balanced match in the distribution of labeled wines across our sample.

The Scanner Data

We use a weekly scanner data set for Northern Californian stores from January 2003 to May 2006. We choose the four control Northern Californian stores (among a total of 38 potential stores to compare to) that match the treated store in terms of pre-period trends in labeled wines. The data provide a unique wine product code identifier (UPC), the name of the wine (including varietal), the number of bottles sold, the pre-discount (gross) shelf price paid, and any retail discount pricing offered. We aggregate the weekly sales data to the month-level for each store to generate the total number of bottles sold per month, the average shelf price, the average price paid (the shelf price net of discounts), and whether a bottle of wine was discounted during the given month. In the treated store, we have 10,508 observations for monthly sales of treated wine products. In the four control stores, the number of observations total to 45,585 with the same average number of observations by each of the control store.

For those wines for which proprietary wine score data exist, we merge the wine score data into the scanner data. In addition, we collect a detailed product attribute data set, identifying the brand of the wine product, varietal (such as Chardonnay, Cabernet, Merlot, Pinot, etc.), type (red, white, or rose), regional designation, and imported status, which we merge with the scanner data set.

Summary statistics are reported in Table 1 for April, the treatment month, and for March, the pre-treatment month in 2006. We report descriptive statistics for the treated store and for the control stores in the first and second columns, respectively. The summary statistics

report average quantity sold during the pre-treatment (March) and during the treatment month (April), along with the standard deviations. In the treated store, 22 bottles per product among the labeled wines are sold on average, while 11 bottles are sold in the control stores during the month of March. Average prices in March and April are 11 dollars for treated wines. The averages are not statistically different between treated and control stores. Approximately 93 percent of the wine consumers purchase is discounted in March and in April, across both the treated and control stores. Average and standard deviation of scores for treated and control stores are similar and average scores are 83 and standard deviation are 3. Additionally, Figure 1 attests that the treated and control stores had similar average scores along with very similar score distributions. In both the treated store and control stores, 50 percent of products are red wines. The proportions of white wine are also similar across the treated and control stores.

Ultimately, there are three kinds of wine products in the sample: the labeled wines, i.e., the treated wines with scores, unlabeled wines with scores (that is, wines that although having been rated with an expert score do not receive the treatment label in our experiment), and the unlabeled wines without scores. Moreover, among labeled (treated) wines it is not always the case that a higher score is associated with a higher average price and or a higher maximum price level displayed.[1]

The treatment store has 20% of the observations in the data, we have one treatment store and four control stores with the same average number of observations by each of the control store. In 2006, the post treatment period consists of a total of number of 748 observations

---

[1]See the online Appendix for summary statistics of the treated wines and comparisons among the three kinds of wines in the sample.

in the post period (months of April and May) and 3059 observations in the controls, (an average of 764 per control store) in the post period. Regarding the pre-treatment period (from January to March), we have the same proportion of observations, the observations for the treatment groups are one fourth of the number of observations of the four controls, with 1160 and 4858 observations respectively (an average per control store is 1215) in 2006.

Given the total quantity Q of wine sold monthly by store, we construct product market shares by dividing each product's quantity sold by the total quantity Q. At most, a wine product represents 8 percent of total monthly wine sales in a store, and the density of market shares are very similar between treated and control stores. To estimate the causal effect of revealing expert scores on consumer demand and valuation, it is crucial that there are similar pre-period trends across treated and control stores for products in the analysis, with respect to quantity sold and market shares. Figure 2 shows that trends in the monthly total market shares of labeled products are quite similar between the treated and control stores. The similarity in trends allows us to investigate the causal effects of the introduction of displaying score labels on demand choices and, given structural demand estimates, to infer consumer WTP for the provision of scores and then perform counterfactual simulations given demand and supply pricing models.

The Structural Consumer Random Utility Demand Model

We model wine demand as a random utility model from buying wine. Using a store-product-month-level panel data set, along with product characteristics data, we estimate flexible specifications of discrete choice structural revealed preference models of consumer demand

(McFadden 1974; Train 2002). Modeling consumer choice as the demand for product bundle of observable attributes, we are able to estimate a dollar value for each attribute. Values of consumers' WTP for expert information are empirically estimated through the addition of an expert opinion attribute to the product space, which is introduced through the field experiment. The discrete choice model (e.g. Berry, Levinsohn and Pakes 1995; McFadden and Train 2000; Nevo 2000; Nevo 2003; Swait and Andrews 2003) also offers flexibility in incorporating heterogeneity in preferences via a random coefficient.[2]

Starting from a random utility framework (e.g. McFadden 1974; Train 2002), where both the product attributes as well as a random term are assumed to enter linearly, the utility from buying a certain product j at time t for the consumer i can be described as:

$$(1) \qquad U_{ijt} = a_j + a_t - \alpha_i p_{jt} + X_{jt}\beta + \gamma T_{jt} Score_j + \xi_{jt} + \varepsilon_{ijt},$$

where a product j is defined as a particular wine UPC sold at a certain store, a_j is a product (UPC-Store) fixed effect capturing the intrinsic preference for product j and a_t is a monthly time fixed effect to control for unobserved demand shocks across periods. The shelf price of product j at month t is denoted by p_{jt} and the marginal utility of price is α_i. In γ, we measure consumers' marginal utility for the labeled score experimentally displayed on

---

[2]When confronting the hedonic (Rosen 1974) and discrete choice models (McFadden 1974), Wong (2018) shows that under certain conditions, an hedonic model and a standard logit model can give similar marginal WTP. Some papers compare WTP derived from both models: they generally found that WTP are higher with discrete choice models. Sinha, Caulkins, and Cropper (2018) estimate both models to get WTP. They think that WTP are better estimated as discrete choice models allows to incorporate more heterogeneity and uses information on market shares, which the hedonic method does not, it only uses prices.

product j. T_{jt} is a dummy variable that is equal to one during the treatment period in the treatment store and equal to zero otherwise, and $Score_j$ is the value of the displayed score for product j. A treated store indicator, a treated wine indicator, a treated month indicator and some interactions are included in X_{jt}. The vector X_{jt} also contains regional, varietal or color information about the product j. The term ξ_{jt} accounts for monthly changes in factors such as shelf space or positioning of the product, among others, that affect consumer utility, and are observed by consumers and firms but not by the researcher. Lastly, ε_{ijt} is an i.i.d. type I extreme value distributed error term, that captures consumer idiosyncratic preferences.

To allow for category expansion or contraction, we include an outside good (no-purchase option), indexed by $j = 0$, whose mean utility is normalized to zero. Therefore, its utility is given by the idiosyncratic term only: $U_{i0t} = \varepsilon_{i0t}$.

Let the α_i coefficient vary according to

$$(2) \qquad \alpha_i = \alpha + \sigma v_i, \ v_i \sim N(0,1),$$

where α and σ, parameters to be estimated, represent the mean utility for price and the deviation to this mean across consumers respectively, and v_i captures unobserved household characteristics. We assume a parametric distribution for v_i, denoted by $P(v)$, which is independently and identically distributed as a standard normal distribution.

We rewrite the utility of consumer i for product j at period t as:

$$(3) \qquad U_{ijt} = \delta_{jt}(p_{jt}, X_{jt}, T_{jt}, Score_j, \xi_{jt}; a, \alpha, \beta, \gamma) + \mu_{ijt}(p_{jt}, v_i; \sigma) + \varepsilon_{ijt},$$

where δ_{jt} is the mean utility and μ_{ijt} is the deviation from the mean utility that allows for consumer heterogeneity in the marginal utility response to the treatment.

The probability that good j is chosen is the probability that good j maximizes the consumer i's utility and results in the predicted probabilities equal to:

$$(4) \qquad S_{ijt} = \frac{exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{N} exp(\delta_{kt} + \mu_{ikt})}$$

where N is the total number of products.

Let the distribution of μ_{ijt} across consumers be denoted by $F(\mu)$. Then the aggregate probability S_{jt} of product j at period t across all consumers is obtained by integrating the consumer level probabilities:

$$(5) \qquad S_{jt} = \int \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{N} \exp(\delta_{kt} + \mu_{ikt})} dF(\mu).$$

Estimating Average Marginal Utility and Price Endogeneity

The demand model represents consumer choice between different wine products over time, where a product is perceived as a bundle of attributes, including expert scores and price. A product-store fixed effect is included to capture constant observed and unobserved product (UPC-Store) factors that affect demand. The econometric error that remains in ξ_{jt} will therefore only include the changes in unobserved product characteristics across time periods such as unobserved consumer level determinants of demand. Prices are set at the wine-price-marketing division level, which covers all of the stores in the sample. Prices are then determined for a set of stores, rather than at an individual store level. Moreover, we

are certain that prices do not adjust during the experiment as explained. However, some endogeneity price issues could remain from omitted brand characteristics that would vary cross time period and we should take into account. In particular, in the presence of endogeneity, the marginal utility of price α would be biased towards zero given that consumers see unobservable features of each wine product ξ_{jt} in (7) when making choices that could be positively correlated with prices. As a consequence, the resulting estimates of WTP would be overestimated, and the welfare effects would also be biased. It is therefore important to use instruments in random utility demand estimation (Ackerberg and Rysman 2005; Bajari and Benkard 2003; Reynaert and Verboven 2014).

In order to address this concern we follow two strategies. First we take advantage of the panel structure of the data and control for unobserved time-invariant determinants of demand with fixed effects. In addition, to take into account for unobservable features that could vary across time periods, we use instruments for prices. Excluded cost shifters, such as input prices that are independent of the wine demand error term, are good instrumental variables but difficult to find. We then instead use instruments that are the characteristics of competing products (Berry, Levinsohn, and Pakes 1995). They are good candidates as they are independent of the utility of the product considered (the demand of the product j depends only on the characteristics of the product j) but they shift equilibrium markups as the economic theory tell us (see Berry, Levinsohn, and Pakes 1995), and then they shift retail prices. In our structural random utility model, we assume that the utility of the consumers depends on the varietal or color information as product characteristics. Considering this information for competing products leads to compute the number of competing products by varietal or color (red, white, or rose). The intuition is that consumer substitution patterns

are stronger within these different types of wines. A product then competes more with products of the same type, and hence, equilibrium markups depend on the presence on competing products within each wine type.

The empirical investigation is presented in Table 2. This Table presents the Logit results from estimation equation (7) as an OLS specification in column 1 and the fixed effects specification in column 2. In the last column, the IV demand specification investigates endogeneity of prices formally and assesses the changes in the marginal utility of price α. To estimate demand parameters controlling for price endogeneity issue, we use a 2SLS estimation method. The dependent variable in all specifications is the $ln(S_{jt}) - ln(S_{0t})$. The covariates in the demand model used are product fixed effects, randomly assigned treatment status by store and month, the score, and price.[3] The first row pertains to the estimate of the marginal utility of price. In the Ordinary Least Squares (OLS) specification, we find that consumers significantly dislike increases in price; the significant estimate of α is -0.075. When we add product fixed effects in column 2, the marginal utility of price becomes more negative, which is the direction away from the omitted variable bias of OLS that was driving α towards zero. The coefficient is now significant and equal to -0.235. When using instrumental variables, the coefficient becomes slightly more negative, now -0.245; the first stage F statistic is large and significantly different from the critical value (F stat is 81.4, with a p value of 0.000), so we reject the null that the first stage coefficients of the two instrumental variables are jointly zero when explaining prices. In practical terms, however, we note that when using

---

[3]It is not the case that all higher scores are correlated with higher prices, and moreover, the variation we use is a score treatment for a subset of wines with scores, so we do not have a high correlation of treatment and score variable with the price variable (cf the online Appendix).

panel fixed effects as controls we deal with most of the omitted variable bias concerns of the OLS specification. In the rest of the paper, we use instruments and fixed effects as controls in order to estimate unbiased marginal utility for price that is then used in the welfare estimates.

Estimating Heterogeneous Marginal Utilities

Given the concern investigated in the simple Logit analysis, we proceed to allow for heterogeneity in demand by adding random coefficients and also instrument for price. We estimate the random parameters Logit demand model from product (UPC-store) monthly market shares using the GMM-estimator proposed by Berry, Levinsohn, and Pakes (1995) and Nevo (2001). We only allow for consumer unobserved heterogeneity in the treatment sensitivity given that we do not have demographic variables in the stores. When estimating demand, the goal is to derive parameter estimates that produce product market shares close to the observed shares. Predicted market shares of each wine product j for draw i are given by

$$(6) \qquad \hat{s}_{ijt} = \frac{\exp(\hat{a}_j + \hat{a}_t - \hat{\alpha}_i p_{jt} + X_{jt}\hat{\beta} + \hat{\gamma}T_{jt}Score_j)}{1 + \sum_{k=1}^{N} \exp(\hat{a}_k + \hat{a}_t - \hat{\alpha}_i p_{kt} + X_{kt}\hat{\beta} + \hat{\gamma}T_{kt}Score_k)}.$$

where the market share is obtained by averaging (6) over all the standard normal draws, that is given by $\hat{s}_{jt} = \frac{1}{R}\sum_{r=1}^{R} \hat{s}_{rjt}$, where R are the number of standard normal draws of v_i.

As shown, this procedure is non-linear in the demand parameters to be estimated. We follow Berry (1994), who constructs a demand side equation that is linear in the parameters to be estimated. This follows from equating the estimated product market shares[4] to the

───────────────

[4]For the random coefficient logit model, the product market share in equation (5) is approximated by the

observed shares and solving for the mean utility across all consumers, defined as:

$$(7) \qquad \delta_{jt}(a, \alpha, \beta, \gamma) = a_j + a_t - \alpha p_{jt} + X_{jt}\beta + \gamma T_{jt} Score_j + \xi_{jt}.$$

For the random coefficient logit model, solving for the mean utility (as in Berry 1994) has to be done numerically (see Berry, Levinsohn, and Pakes 1995; and Nevo 2001). Once this inversion has been made, one obtains equation (7) which is linear in the parameter associated with all wine attributes. The estimates are obtained by a fixed effects OLS regression. We let a, $\alpha, \beta, \gamma, \sigma$ be the demand side parameters to be estimated, where the linear parameters are $(a, \alpha, \beta, \gamma)$ and σ is the price non-linear random coefficient parameter. In the random coefficient logit model, the parameters are obtained by feasible Simulated Method of Moments (SMOM) following Nevo's (2000) estimation algorithm, which requires equation (7).[5] Given the estimated demand marginal utility parameters, we finally obtain estimates of average WTP for the labels by dividing the estimates for γ by the average marginal utility of price α.

The demand model estimates are presented in Table 3. In column (1), we estimate demand with respect to a regional dummy for California, a discounted price, a score, treatment store and period interactions with and without the score variable, and the coefficients of interest, associated with the "Score X Labeled Wines X Treated Store X Treated Period" and

---

Logit smoothed accept-reject simulator.

[5]The aim is to concentrate the simulated GMM objective function such that it will be only a function of the non-linear parameters. By expressing the optimal vector of linear parameters as a function of the non-linear parameters and then substituting back into the objective function, it can be optimized with respect to the non-linear parameters alone.

"Labeled Wines X Treated Store X Treated Period" variables. In column (2), we estimate demand as in column (1), with the inclusion of varietal fixed effects. Finally, column (3) repeats the specification in column (2) but uses instrumental variables for price. Starting with the price coefficients, in column (1), in an OLS specification, the estimated average marginal utility of price is −0.078 and is significant. When adding varietal fixed effects, the price coefficient becomes more negative −0.166, and the instrumental variables estimate in column (3) for the marginal utility of price is −0.232. We also find significant price hetero-geneity given the standard deviation equal to 0.1. The preferred specification is in column (3), as it was also for the Logit specifications, using instrumental variables to deal with the endogeneity of prices when estimating demand.

We also feature all the lower order terms of the interactions of interest that appear in Table 3. For instance, given the non-significant coefficient of "Treated Store X Treated Period", we see that unlabeled wines do not have a significant mean utility. For the two difference in difference causal coefficients of interest, we find that the "Labeled Wine X Treated Store X Treated Period" has a negative mean utility of −0.498 and that the mean utility of each score point labeled, given by "Score Level X Label Wine X Treated Store X Treated Period", is significant and equal to 0.007. Taken together, we interpret the joint estimate given as $-0.498 + 0.007 \text{ x } ScoreLevel$ as the causal effect on the mean utility of our treatment. This causal effect is estimated to be a negative mean utility for scores less than $0.498/0.007 = 71.4$ and positive for scores larger than 71.4. Looking at the "Score Level X Label Wine X Treated Store X Treated Period" coefficient in Table 3 , namely 0.007, we once again reject the null that the score labels just produce a "hype" effect on demand, common to all labeled products willingness to pay. In fact, our evidence is consistent with

label content specific information, in terms of the score level, significantly affecting the mean utility of consumers.

Finally, in terms of wine regions, consumers place a positive and significant marginal utility for California wines and for most varietals.

WTP for Wine Attributes and Expert Opinion Experimentally Displayed Scores

Given the estimated model parameters in column (3) of Table 2, and also in addition column (3) of Table 3, we first recover, for each score level, the implied WTP for the labels. The top Panel A of Figure 3 displays the predicted WTP on the vertical axis, for score levels on the horizontal axis, by dividing the marginal utility of labels and scores in the treated store and the treatment period by the average marginal utility of price following Train (2003). We then obtain estimates of the average WTP for a change in a score point. The solid line features the predicted additional average WTP for labels, given the Logit estimates in column (3) of Table 2, and is a linear function of the score given by $\frac{-0.368+0.005*Score}{0.245}$. The dashed line represents the average WTP given the Random Coefficient Logit estimates of column (3) in Table 3, and is given by $\frac{-0.498+0.007*Score}{0.232}$. First, our demand estimates suggest that consumers would need to be offered a discount if wines feature a score less and equal than 71, given that, for the Logit model, when the score is below 72 the WTP becomes negative, whereas for the random coefficient this happens for scores below 71. Second, on average, consumers value a score point increase by about 0.5 cents, which corresponds to the slope of the solid line of the Logit specification, to 0.7 cents for the random Coefficient specification in the dashed line. As a consequence, for a bottle featuring an average score of 83 consumers would be willing to pay additionally between 20 cents (solid line) and 60 cents

more (dashed line) if the score was revealed to them in the label.[6]

In the bottom Panel B of Figure 3, we depict for each of the labeled score levels in the data on the horizontal axis, the average price per score, and the total estimated average WTP of the wines using Logit and Random coefficient Logit demand estimates. For both demand estimates, we see that the WTP is similar to the average price for scores of 72 and that consumers are willing to pay more than the average price for all other score levels.

Welfare Changes in Counterfactuals With Expert Scores' Labels

In order to investigate the welfare effects due to expert opinion shelf labels being implemented, we first specify a supply model and consider three cases for the price decisions of suppliers. First, we assume that there is perfect competition, prices are set equal to marginal costs, and we recover the highest bound on marginal costs. Therefore, we can estimate changes in welfare with and without shelf labels when prices do not adjust given counterfactual labeling scenarios. Then, we consider the case when suppliers have market power. We can also provide bounds on marginal costs of products under noncompetitive pricing behavior. First, assuming that prices are chosen in a Bertrand Nash fashion, we recover an intermediate level of marginal costs by subtracting from prices the margins under Bertrand Nash. Second, assuming monopoly pricing, we recover the lower bound on products' marginal costs, given by the difference between observed prices and implied monopoly (larger) markups.

---

[6]Assessing the WTP of expert opinions in the wine market is a previously unexplored avenue in the literature. However, we could find some WTP estimates for some other characteristics of wine, deduced from consumer questionnaires (e. g. Loureiro 2003; Sellers-Rubio and Nicolau-Gonzalbez 2016).

Given all these supply scenarios, we simulate counterfactual choices and new equilibrium prices, and estimate bounds on the resulting implied welfare changes. In particular, we break up the effect of expert opinion score labels on welfare into two effects. Keeping prices constant, which corresponds to a competitive supplier model, we estimate the pure consumption switching effect with provided expert opinion score labels. Allowing for prices to adjust, we estimate the competitive change effect in the counterfactual simulations with expert score labels.

Counterfactual Simulations under Competitive Supplier Case

Given that marginal costs are constant regardless of expert opinion score shelf labels implemented and regardless of the score levels provided, prices of products remain unchanged when we simulate consumer choices with expert opinion score shelf labels. We thus focus on the changes in consumer surplus only in order to estimate the welfare effects of counterfactuals assuming a competitive supply side. Over the pre treatment period where we observe the consumer choices given no shelf labels, we simulate consumer choices with labels keeping prices unchanged. We then compute, in both situations when the shelf labels are available or not, the consumer surplus, ceteris paribus (prices unchanged) following Small and Rosen (1981) and also the approach found in Allenby, et. al (2014). The expected consumer surplus, CS_i, which corresponds to the compensating variation for a change in product attributes (Small and Rosen 1981), can therefore be defined as:

$$(8) \qquad CS_i = \frac{1}{|\alpha|} ln \sum_j e^{a_j + a_t + X_{jt}\beta_x - \alpha price_{jt}} + C,$$

where α denotes the average marginal utility of price and C is a constant. We then compute the consumer surplus when shelf labels are not displayed, the consumer surplus when shelf labels are displayed and given the label claims. We then obtain the estimated changes in consumer surplus from introducing the expert score labels in the treated stores as follows, where the constant C will be differenced out:

$$(9) \qquad \Delta CS = \sum_i \sum_t [CS_{i,Labels} - CS_{i,NoLabels}].$$

CS_i is given by (8) and ΔCS denotes the total CS with labels minus the total CS without labels.

In our case, we change the format in which the quality information is presented to the consumer in the form of a score. Any difference in consumer surplus, therefore, reflects only a possible unobserved change in consumer perceptions of attributes rather than underlying actual quality differences (Nevo 2003) and allows us to interpret these changes as causal effects of the additionally posted expert opinion scores on shelf labels, and consumer valuation of this additional information.

We use a long historical pre treatment period, from January 2005 to March 2006, and the structural demand estimates to simulate what consumer's choices would be in the presence of labels and to estimate change in consumer surplus from choices made when labels are not available. The difference in surplus amounts to the welfare change due to the labels when prices do not change. In other words, a positive change in surplus when adding labels means that they are significantly valued by the consumer. By adding the total change in consumer surplus, we obtain a total value of displaying the labels in this particular market assuming

constant prices.

Counterfactual Simulations in a Noncompetitive Supply Side Model Case

In this case, we consider the retailer choosing prices to maximize profits. First, we assume a Bertrand Nash competition, that is the retailer maximizes the profits from selling each product given the price of the other products sold in the retailer. We also consider the most extreme case scenario of market power in which prices for all products are set to maximize the retailers' total profit. Under these supply side scenarios, prices will change if labels are present, and we will simulate the consumer choices, new equilibrium prices and resulting welfare changes if prices adjust in the Bertrand Nash competition case or in the monopoly multi-product retailer case.

The approach is as follows. First, given estimated demand, we recover marginal costs by subtracting from observed prices the implied markups of the two supply cases considered. In so doing, we obtain upper and lower bounds on marginal costs for all of the products, by assuming both assumptions about the strategic behavior of the retailer described above. The last step is the counterfactual simulation approach of adding expert opinion labels and solving for new equilibrium prices in the pre treatment period under both supply scenarios.

Recovering Marginal Costs under Bertrand Nash Pricing

Every producer is maximizing its profits $\pi_j = (p_j - c_j)s_j(p)$ by choosing each price p_j given the competitor prices and marginal costs c_j, where p is the vector of retail prices.[7]

---

[7]The producer maximizes its profit at each time period t. We remove the subscript t from notations for simplicity reasons.

Assuming Bertrand Nash pricing we have the first order conditions: $\forall j = 1, ..., N, \frac{\partial \pi_j}{\partial p_j} = s_j(p) + (p_j - c_j)\frac{\partial s_j}{\partial p_j}$.

Let's define a matrix Δ as

$$\Delta(p) = \begin{bmatrix} \frac{\partial s_1(p)}{\partial p_1} & \frac{\partial s_1(p)}{\partial p_2} & \cdots & \frac{\partial s_1(p)}{\partial p_{15}} \\ \frac{\partial s_2(p)}{\partial p_1} & \frac{\partial s_2(p)}{\partial p_2} & \cdots & \frac{\partial s_2(p)}{\partial p_N} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial s_N(p)}{\partial p_1} & \frac{\partial s_N(p)}{\partial p_2} & \cdots & \frac{\partial s_N(p)}{\partial p_N} \end{bmatrix}$$

which is composed of the derivatives of market shares with respect to prices, evaluated at the average marginal utility of price. Let T^{bn} be the ownership matrix of the producer, which corresponds to the identity matrix in this case. We can solve for Bertrand-Nash price cost markups as $m^{bn} = p - c^{bn} = -(T^{bn}\Delta(p))^{-1}s(p)$ where c^{bn} is the vector of Bertrand Nash marginal costs and s(p) is the vector of market shares.

The upper bound on marginal costs c^{bn} given Bertrand Nash is then

$$\hat{c}^{bn} = p - m^{bn} = p + (T^{bn}\Delta(p))^{-1}s(p).$$

Recovering Marginal Costs under Monopoly Pricing

In this case, the prices are chosen to maximize the joint profit $\pi = \sum_{k=1}^{N}(p_k - c_k)s_k(p)$, and the first order conditions are: $\forall j = 1, ..., N, \frac{\partial \pi}{\partial p_j} = s_j(p) + \sum_{k=1}^{N}(p_k - c_k)\frac{\partial s_k}{\partial p_j}$.

Given the Δ matrix and letting T^m be the ownership matrix for joint profit price choices where all elements are 1, then we can solve for monopolist price cost markups as $m^m = p - c^m = -(T^m \times \Delta(p))^{-1}s(p)$ where c^m is the vector of monopolist marginal cost and $\times$ is

an element by element multiplication.

We then obtain the lower bound of marginal costs c^m given monopolist pricing as

$$\hat{c}^m = p - m^m = p + (T^m \times \Delta(p))^{-1} s(p).$$

Counterfactual Simulations and Estimated Welfare Changes

The final step is to perform simulations for estimated welfare changes under expert opinion label counterfactuals, denoted as *Label*. In the pre treatment period from January 2005 to March 2006, we allow consumers to face the expert opinion labels when they make their choice under several supply scenarios given the costs recovered, and simulate the new equilibrium prices and quantities in these counterfactuals. In particular, we simulate the equilibrium $(N \times 1)$ vector of retail prices p^* under displayed shelf labeling scenarios *Label* and assuming that retailers follow either a Bertrand Nash pricing game:

$$(10) \qquad p^* = \hat{c}^{bn} - \left(T_r^{bn} * \Delta_r(p^*)\right)^{-1} s(p^*, Label),$$

or assuming that retailers follow a monopoly game:

$$(11) \qquad p^* = \hat{c}^m - \left(T_r^m * \Delta_r(p^*)\right)^{-1} s(p^*, Label).$$

We assess the changes in the welfare components (consumers' and retailers' surplus) resulting from the changes of the simulated counterfactual equilibrium prices p^* under the labeling and different supply side scenarios, relative to the baseline observed equilibrium

prices p in the pre period, without the labels. The expected consumer surplus, CS_i is defined as (8). We then obtain the estimated changes in consumer surplus from adding the expert score shelf labels, as given by (9), which is defined as the total consumer surplus with labels minus the total consumer surplus without labels. Given the new prices and market shares, we estimate the changes in retail surplus as the change in the profits, and the sum of that and changes in consumer surplus results in the total welfare changes. The goal is to compare the estimated welfare effects of alternative labeling counterfactuals varying the degree of ability of retailers to mark-up over the costs.

Simulation Results and Estimated Welfare Changes due to Experimental Shelf Labels

Given the structural demand and supply models, demand estimated parameters and estimated costs, we use the pre-period in the treated store from January 2005 to March 2006 and simulate what would be the choices and resulting equilibrium prices if products received the expert opinion score labels. In each simulation, we consider three supply settings: a competitive one where prices do not change, a Bertrand Nash one, and a monopoly setting. By adding expert opinion labels and resulting simulation, we interpret the effect given constant prices (competitive supply) to be due to "Pure Consumption switching." In addition, by allowing prices to adjust as Bertrand Nash or as a local monopolist, this allows us to estimate the "Competition effect" of changing the labels, given that firms may adjust the prices when scores are displayed and introduce product differentiation into the choice set.

The top panel of this Figure 4 shows the average prices at baseline and corresponding 95% confidence interval, and next to that the simulated average price and 95% confidence interval for simulated prices of adding labels under Bertrand Nash pricing competition.

We see that we cannot reject that prices under Bertrand Nash with labels are similar to the average baseline prices. In contrast, looking at the average price from simulating a scenario post labels under monopolist supply, we see a significant increase in average price. Not only does average price increase, the entire distribution of prices post labels under monopoly pricing would shift to the right of the baseline price distribution, as we see in the bottom panel of Figure 4. The bottom panel has the estimated kernel density of the price distributions under baseline no Labels (black line density), Bertrand Nash Case (grey line density - overlapping black one) and Monopolist Case with Labels (dashed line density). We see that the monopolist case density is to the right of the baseline and Bertrand Nash densities, which means prices overall went up with labels if firms have monopolist market power, but prices remain basically unchanged with labels if firms are competing as Bertrand Nash.

In Table 4, we present changes resulting from adding expert opinion score shelf labels under competitive (row 1), Bertrand Nash (row 2), and monopoly pricing (row 3). Then, in this table's columns, we report first changes in average prices, then average change in consumer surplus, average change in profits, and average change in total welfare. All are in dollars per consumer (per capita). Standard errors are reported in parentheses. They are then to be interpreted as the value of the shelf labels to consumers, retailer and welfare overall.

We see that adding expert opinion score labels has a significant positive effect on consumer surplus and welfare if prices remain constant, under the competitive supply case. If firms are allowed to adjust prices post shelf labels, and are competing as Bertrand Nash, prices increase by a very small amount, consumer surplus drop as compared to the competitive case,

albeit in an economic insignificant way. Relative to baseline, by introducing scores in the form of labels, this increases significantly retailer surplus and welfare in the Bertrand Nash supply case. In this scenario, average prices increase significantly by less than one cent (an economically minuscule price effect), consumer surplus increases by 2.02 dollars per capita, and welfare increases by 2.38 dollars per capita, while profits increase significantly by 36 cents per capita. Finally, in the monopolist extreme case, by introducing scores in the form of labels, this increases significantly consumer surplus, retail surplus and welfare relative to baseline. In this scenario, average prices increase significantly by 45 cents, consumer surplus increases by 1.59 dollars per capita (a significantly smaller increase than in the two other supply competitive cases), and welfare increases significantly by 2.21 dollars per capita (which is less than the Bertrand Nash case), while profits increase significantly by 64 cents per capita (almost twice than in the Bertrand Nash Case). This would imply that 30% (15% in the Bertrand Nash Case) of the per capita added value would go to the retailer. Our average price increase is quite small when compared to the price effects of 2.3 Euros, estimated by Ali et al (2014) for wines in the Bordeaux region. One possible reason is that the sample of wines in their data are quite different from ours, have higher quality and also of higher prices, given that our range of scores is 72 to 90, and their sample does not have such low scores and features several scores in the mid 90s.

Conclusion

This paper estimates flexible discrete choice demand and supply models to infer whether consumers place a significant value on the reduction of asymmetric information about wine

quality, in the form of expert opinion score labels. We use a product store month data set for treated and control stores, consisting of retail scanner data at the time of a field experiment that reveals the expert opinion scores to retail consumers at the point of purchase, to estimate the demand and supply structural model primitives, namely demand parameters and marginal costs. In particular, we obtain an implied average willingness to pay of 0.5 to 0.7 cents per score point displayed in the labels and we find that there is heterogeneity in WTP for wines originating from different regions.

Using counterfactual simulations, we estimate the changes in consumer surplus, retail profits and welfare due to available quality information from expert opinion score labels, assuming different benchmarks of retailer pricing behavior given the estimated demand model and recovered marginal costs under these different supply models.

We show that adding expert opinion shelf labels increases differentiation and allows prices to increase and firms to compete less, leading to an overall significant welfare improvement given that (i) consumers significantly value the score information, and (ii) the profits increase when firms are able to incorporate the new shelf labels when setting prices as Bertrand Nash (and more so in a coordinated monopolist fashion). We estimate there to be a significant gain in welfare of 2 dollars per capita, which when considering there to be about 10,000 consumers by month (based on auxiliary data on a store at the consumer level) to be about 20,000 dollars, representing 7.8 percent of the revenues in this market. This suggests that disclosing expert opinions results in small but significant positive welfare effects. Extrapolating to the national market, given total US wine retail revenues for 2013 were \$36.3 billion dollars,[8] our findings would imply that welfare would increase by 2 billion dollars from disclosing expert

---

[8]http://www.wineinstitute.org/resources/pressroom/04242014.

opinions about the quality of wines in the form of the experiment labels.

References

Ackerberg, Daniel A. 2001. "Empirically Distinguishing Informative and Prestige Effects of Advertising." *The RAND Journal of Economics* 32(2): 316-333.

Ackerberg, Daniel A. 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination." *International Economic Review* 44 (3):1007-1040.

Ackerberg, Daniel A. and Mark Rysman. 2005. "Unobserved Product Differentiation in Discrete-choice Models: Estimating Price Elasticities and Welfare Effects." *RAND Journal of Economics* 36 (4): 1–000.

Allenby, Greg, Jeff Brazell, John Howell, and Peter Rossi. 2014. "Valuation of Patented Product Features." *The Journal of Law & Economics* 57(3): 629-663.

Ali, Héla Hadj, Sébastien Lecocq, and Michael Visser. 2008. "The Impact of Gurus: Parker Grades and En Primeur Wine Prices." *The Economic Journal* 118 (June): 158–173.

Anderson, Michael and Jeremy Magruder. 2012. "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database." *Economic Journal* 122 (563):957-989.

Bajari, Patrick and C. Lanier Benkard. 2003. "Discrete Choice Models as Structural Models of Demand: Some Economic Implications of Common Approaches." *Working paper.* https://web.stanford.edu/ lanierb/research/al3-13-03.pdf.

Berger, Jonah, Alan T. Sorensen, and Scott J. Rasmussen. 2010. "Positive Effects of Negative Publicity: When Negative Reviews Increase Sales." *Marketing Science* 29(5): 815-827.

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

Berry, Steven, J. Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63:841-890.

Bonnet, Celine, James Hilger, and Sofia B. Villas Boas. 2020. "Reduced Form Evidence on Belief Updating under Asymmetric Information - The Case of Wine Expert Opinions." *European Review of Agricultural Economics.* jbz054. https://doi.org/10.1093/erae/jbz054.

Carson, Richard, Nicholas Flores, Kerry M. Martin, and Jennifer L. Wright. 1996. "Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods." *Land Economics* 72(1): 80-99.

Hensher, David A., Jordan J. Louviere, and Joffre D. Swait. 1998. "Combining Sources of Preference Data." *Journal of Econometrics* 89 (1–2): 197–221

Hilger, James, Greg Rafert, and Sofia B. Villas Boas. 2011. "Expert Opinion and the Demand for Experience Goods: An Experimental Approach." *Review of Economics and Statistics* 93 (4): 1289-1296.

Jacobsen, Grant D. 2015. "Consumers, Experts, and Online Product Evaluations: Evidence from the Brewing Industry." *Journal of Public Economics* 126:114-123.

Jin, Ginger and Phillip Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *The Quarterly Journal of Economics* 118(2): 409-451.

Johnson, Justin P. and David P. Myatt. 2006. "On the Simple Economics of Advertising, Marketing, and Product Design." *American Economic Review* 96 (3): 756-784.

Kiesel, Kristin and Sofia 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): Art 4.

Kim, Donghun. 2004. "Estimation of the effects of New Brands on Incumbent's Profits and Consumer Welfare: The U.S. Processed Cheese Market Case." *Review of Industrial Organization* 25:275-293.

Loureiro, Maria L. 2003. "Rethinking New Wines: Implications of Local and Environmentally Friendly Labels." *Food Policy* 28(5-6): 547-560.

McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand.' *Journal of Public Economics* 3:303-28.

McFadden, Daniel. "Structural Discrete Probability Models Derived from Theories of Choice." *Structural Analysis of Discrete Data and Econometric Applications.* edited by C. Manski and D. McFadden, 198-272. Cambridge: MIT Press, 1981.

McFadden, Daniel. "Computing Willingness-to-Pay in Random Utility Models." Trade, Theory, and Econometrics: Essays in Honor of John S. Chipman. J. Moore, R. Riezman, and J. Melvin (eds.), London: Routledge, 1999.

McFadden, Daniel and Kenneth E. Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15:447-470.

Montgomery, Cynthia and Birger Wernerfelt. 1992. "Risk Reduction and Umbrella Branding." *The Journal of Business* 65(1): 31-50.

Nevo, Aviv. 2000. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." *Journal of Economics & Management Strategy* 9:513-548.

Nevo, Aviv. 2003. "New Products, Quality Changes and Welfare Measures Computed from Estimated Demand Systems." *Review of Economics and Statistics* 85:266-275.

Rao, Akshay R. 2005. "The Quality of Price as a Quality Cue." *Journal of Marketing Research* XLII : 401-405

Reinstein, David A. and Christopher M. Snyder. 2005. "The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics." *Journal of Industrial Economics* 53: 27– 51.

Reynaert, Mathias and Frank Verboven. 2014. "Improving the Performance of Random Coefficients Demand Models: the Role of Optimal Instruments." *Journal of Econometrics* 179: 83-98.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82(1): 34-55.

Sellers-Rubio, Ricardo and Juan Luis Nicolau-Gonzalbez. 2016. "Estimating the Willingness to Pay for a Sustainable Wine using a Heckit Model."*Wine Economics and Policy* 5: 96-104.

Shiv, Baba, Ziv Carmon, and Dan Ariely. 2005. "Placebo Effects of Marketing Actions: Consumers May Get What They Pay For." *Journal of Marketing Research* XLII : 383-393.

Sinha, Paramita, Martha L. Caulkins, and Maureen L. Cropper. 2018. "Do Discrete Choice Approaches to Valuing Urban Amenities Yield Different Results than Hedonic Models?" *NBER Working Paper* No. w24290.

Small, Kenneth A. and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49:105-130.

Swait, Joffre and Rick L. Andrews (2003). "Enriching Scanner Panel Models with Choice Experiments." *Marketing Science* 22(4): 442-460.

Train, Kenneth E. "Discrete Choice Methods with Simulation." Cambridge: University

Press, 2002.

Whitehead, John C., Subhrendu K. Pattanayak, George L. Van Houtven, and Brett R. Gelso, 2008. "Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: An Assessment of the Atate of the Science." *Journal of Economic Surveys* 22: 872–908.

Wong, Maisy. 2018. "A Tractable Approach to Compare the Hedonic and Discrete Choice Frameworks." *Journal of Housing Economics* 41(C): 135-141.

Figure 1: Kernel Density of Scores of Treated Wines in Treated and Control Stores

Note: This figure displays jointly the kernel density estimates of the score distribution for the set of treated products in the treated store and the kernel density estimates of the score distribution in control stores for the same group of wine products treated in the treated store, given that we can see the same products in the control stores. The Kolmogorov-Smirnov (KS) test cannot reject the equality of treated wines scores' distributions in the treated and in the control stores, given that the KS test statistic is 0.0232 (p value 1.000).

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

	(1) Treated Store Wines	(2) Control Stores Wines
Quantity (March)	21.68	10.76
	(0.00)	(0.00)
Quantity (April)	18.16	10.70
	(0.00)	(0.00)
Price (March)	10.51	10.76
	(0.51)	(0.52)
Price (April)	10.45	10.70
	(0.49)	(0.52)
% discounted (March)	93	93
% discounted (April)	92	92
Score	83.53	83.53
	(3.06)	(3.13)
Type		
% Red	49	50
% White	44	42
% Rose	7	7
Varietals		
% Cabernet	15	16
% Chardonnay	26	24
% Gewurztraminer	5	5
% Merlot	16	15
% Pinot Noir	7	7
% Shirah	8	9
% Zinfandel	7	7
% Other Red Varietals	3	3
% Other White Varietals	12	13
Number Wines	404	421
Number Observations	10508	45585

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

Source: Scanner data set. Quantity is measured in number of bottles.

Figure 2: Trends of Market Shares of Treated Wines in Treated and Control Stores
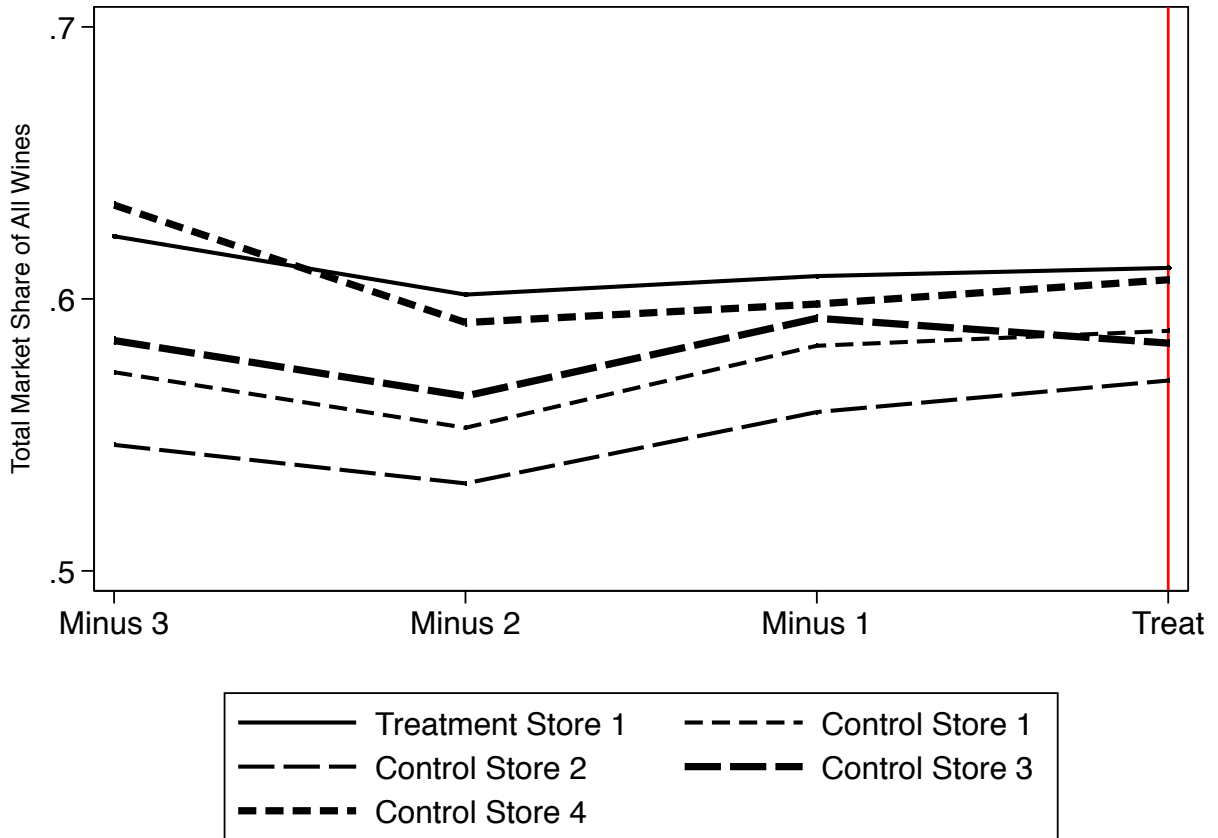
Note: This figure displays jointly the evolution of the treated wine product market shares in the treated and control stores in the pre treatment periods. Minus 1 means one period before the treatment, Minus 2 means 2 periods before, and Minus 3 means 3 periods before.

Table 2: OLS and IV Logit Wine Demand Estimates

	(1)	(2)	(3)
Price	-0.075***	-0.235***	-0.245***
	(0.001)	(0.002)	(0.003)
Treated Store X Treated Period	0.044	0.066	0.065
	(0.115)	(0.076)	(0.076)
Treated Store X Pre Period	0.238***	0.184***	0.183***
	(0.027)	(0.018)	(0.018)
Control Stores X Treated Period	0.124***	0.041	0.039
	(0.048)	(0.033)	(0.033)
Labeled Wines X Treated Period	-0.055	-0.042	-0.040
	(0.052)	(0.036)	(0.036)
Labeled Wines X Treated Store	-0.132***	-0.069***	-0.069***
	(0.029)	(0.021)	(0.021)
Score Level	0.007***	0.000	0.181
	(0.000)	(.)	(0.156)
Score Level X Treated Store X Pre Period	-0.001	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Labeled Wine X Treated Store X Treated Period	-0.335	-0.371**	-0.368**
	(0.250)	(0.168)	(0.167)
Score Level X Labeled Wine X Treated Store X Treated Period	0.006*	0.006**	0.005**
	(0.003)	(0.002)	(0.002)
Constant	-6.117***	-3.926 ***	-20.266
	(0.015)	(0.275)	(13.437)
Num of Obs.	56093	56093	56093
R squared	0.147	0.633	0.632
Product FE	No	Yes	Yes
Instruments for Price	No	No	Yes
First Stage F(2, 56093)			81.400
p value Prob greater than F			0.000

Clustered errors in parentheses at the month level. Controls are best matched stores. The dependent variable is the ln(market share of product)-ln(share of outside option).

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

c

Table 3: Random Coefficient Logit Wine Demand Estimates

	(1)	(2)	(3)
Constant	-6.432***	-6.548***	-7.024***
	(0.017)	(0.017)	(0.017)
Treated Store X Treated Period	0.068	0.095	0.057
	(0.113)	(0.110)	(0.112)
Treated Store X Pre Period	0.236***	0.210***	0.157***
	(0.027)	(0.026)	(0.027)
Control Stores X Treated Period	0.110**	0.137***	-0.007
	(0.047)	(0.046)	(0.047)
Score	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)
Labeled Wine X Treated Store X Treated Period	-0.464*	-0.492**	-0.498**
	(0.247)	(0.240)	(0.244)
Score Level X Labeled Wine X Treated Store X Treated Period	0.008**	0.006**	0.007**
	(0.003)	(0.003)	(0.003)
Price	-0.078***	-0.166***	-0.232***
	(0.001)	(0.015)	(0.025)
California	0.473***	0.419***	0.426***
	(0.013)	(0.014)	(0.015)
Cabernet		0.854***	0.733***
		(0.112)	(0.120)
Chardonnay		1.332***	1.227***
		(0.123)	(0.130)
Gewurztraminer		0.626***	0.486***
		(0.121)	(0.127)
Merlot		0.912***	0.791***
		(0.117)	(0.124)
Pinot Noir		0.997***	0.881***
		(0.131)	(0.133)
Shirah		0.619***	0.486***
		(0.124)	(0.129)
Zinfandel		0.758***	0.637***
		(0.123)	(0.129)
Other Red Varietals		0.747***	0.614***
		(0.131)	(0.135)
Other White Varietals		0.590***	0.499***
		(0.173)	(0.172)
Price SD	0.000	0.100***	0.100***
	(0.009)	(0.011)	(0.009)
GMM			
Num of Obs.	56093	56093	56093
Instruments for Price	No	No	Yes

Clustered errors in parentheses at the month level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Coefficients on Lower Order Interactions of Treatment Store, Period,Score, and Labeled omitted due to space.

Figure 3: Average WTP for Score Disclosure, Average Prices, and Average WTP by Score
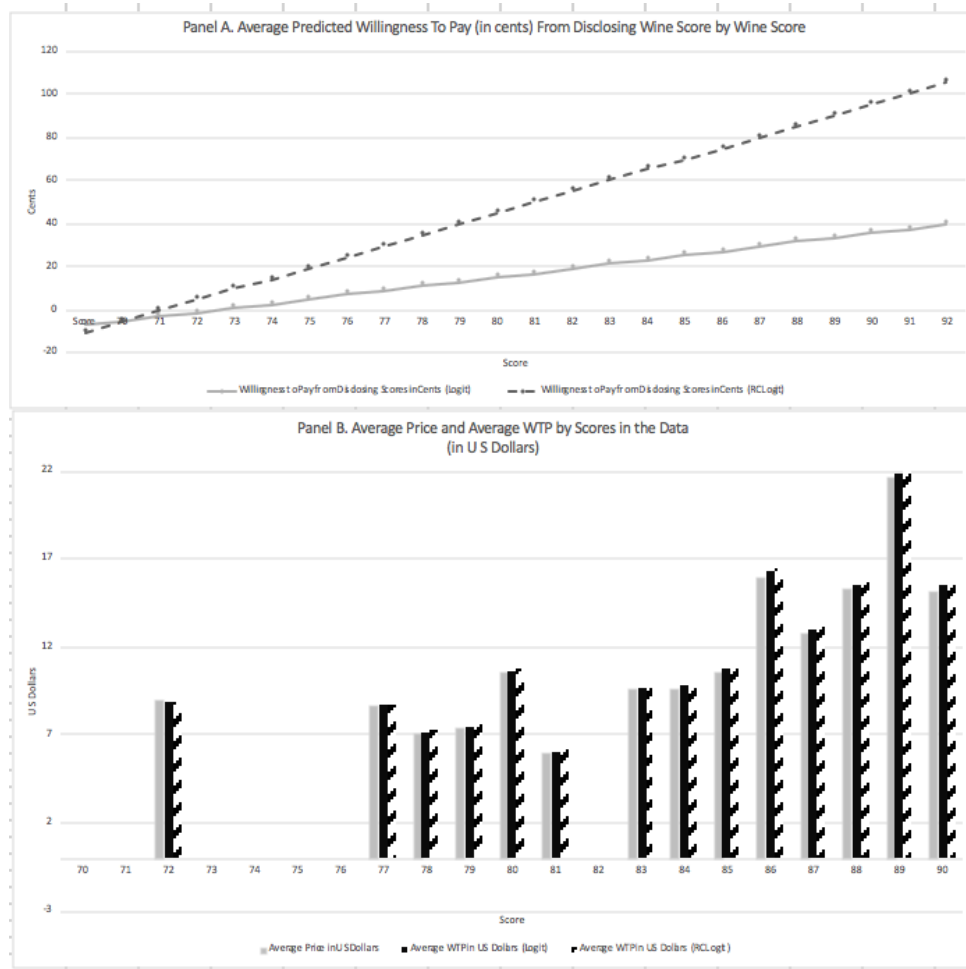
Note: The top panel A of this Figure reports predicted average WTP from displaying the expert score label for each score level: based on the Logit estimates of column (3) in Table 2 in solid line, and in a dashed line based on the RC Logit estimates of column (3) in Table 3. The Bottom panel B depicts average prices (in light grey), and the average WTP for each of the scores in the wine data set: based on the Logit estimates in column (3) of Table 2 in black, and in diagonal pattern based on the random coefficients demand estimates in Column (3) of Table 3.

Figure 4: Baseline Prices and Simulated Prices With Labels under Different Supply Cases
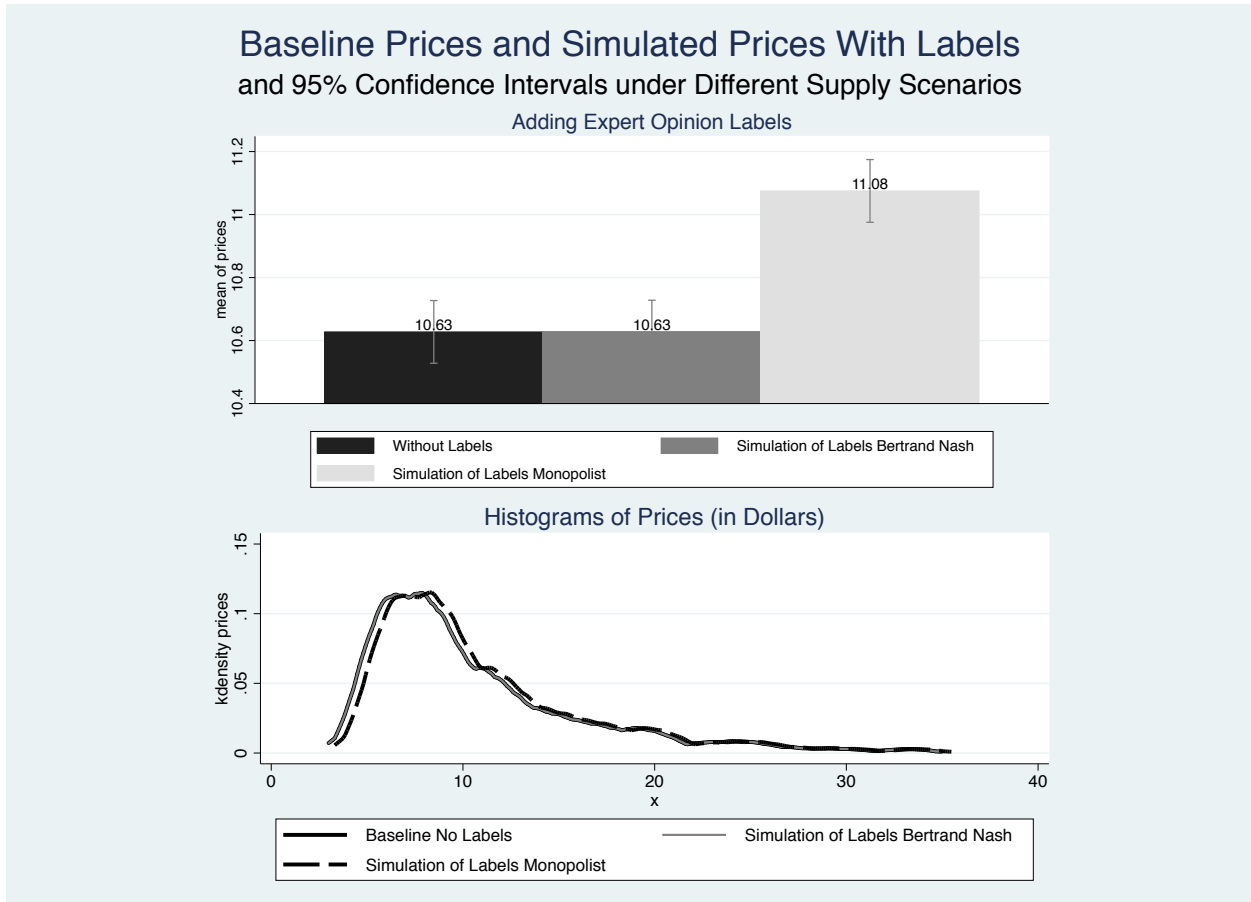
Note: The top panel of this figure reports average prices and 95% Confidence interval for baseline and simulated prices of adding labels under Bertrand Nash and under monopolist supply cases. The bottom panel has the estimated kernel density of the price distributions under baseline no labels (black line density), Bertrand Nash case with labels (grey line density- overlapping black one: no change) and monopolist case with labels (dashed line density). We see that the monopolist case density is to the right of the baseline and Bertrand Nash densities, which means prices overall went up with labels if firms have monopolist market power, but prices remain basically unchanged with labels if firms are competing as Bertrand Nash.

Table 4: Simulated Changes due to Expert Scores' Labels

	Price Change	Average CS Change ($ per capita)	Average Profit Change ($ per capita)	Average Welfare Change ($ per capita)
Adding Expert Opinion Scores				
Competitive Supply		2.027	0.000	2.027
		(0.007)	(0.000)	(0.007)
Bertran Nash Supply	0.001	2.024	0.359	2.383
	(0.000)	(0.007)	(0.005)	(0.011)
Monopoly Supply	0.452	1.578	0.643	2.221
	(0.011)	(0.004)	(0.014)	(0.012)

Simulated Average Changes, and standard errors in parentheses, in Prices, Consumer Surplus, Profits, and Welfare when wines are identified according to expert opinion Scores' labels. All estimates based on the demand estimates and Simulations using the treated store only during the pre treatment period.