Title
Expert Opinion and the Demand for Experience Goods: An Experimental Approach

Journal
Review of Economics and Statistics, 93(4)

Authors
Villas-Boas, Sofia B
Rafert, Greg
Hilger, JAmes

Publication Date
2011

Peer reviewed

Expert Opinion and the Demand for Experience Goods: An Experimental Approach in the Retail Wine Market

James Hilger, Greg Rafert, Sofia Villas-Boas*

January 28, 2010

Abstract

The effect of expert opinion on consumer demand for experience goods is difficult to quantify, as the relationship between purchases and reviews may be driven by product quality. Further, it is unclear whether a review-based demand effect is due to the provision of quality or existence information. Utilizing a field experiment in the retail wine market to overcome these obstacles, we find that there is a significant positive average consumer response to expert opinion labels for wine. We find that while demand decreases for wines that receive low scores, demand for average- and higher-than-average-scored wines increases. The results indicate that expert opinion labels transmit quality information that affects demand as opposed to solely increasing the wine's shelf visibility to the consumer.

*Hilger: U.S. National Marine Fisheries Service; Rafert: Analysis Group; and Villas-Boas: Department of Agricultural and Resource Economics, University of California at Berkeley. We thank the editor Julio Rotemberg, two anonymous referees, as well as Max Auffhammer, Peter Berck, Stefano DellaVigna, Christopher Gustafson, Daniel Hosken, John List, Jeremy Magruder, Doug Miller, Carlo Prato, David Schmidt, Stephen Stohs, Elizabeth Szyleyko, Miguel Villas-Boas, seminar participants at the University of California at Davis, the 35th EARIE Conference in Toulouse, and the FTC & Northwestern University Microeconomics Conference for helpful comments. We are also thankful to Michelle Scizak, Reed Johnson, and Steve Flaming for help with the data; Will Liu and Jeff Coronado for help with the experimental implementation; Kyle Birchard, Elizabeta Perova, Grant Chen, Patricia Javier, Katan Patel, and Elizabeth Creed for their excellent research assistance. The authors gratefully acknowledge the financial support received from the Giannini Foundation. The results are not necessarily those of the U.S. National Marine Fisheries Service. Corresponding author: J. Hilger at james.hilger@noaa.gov.

1 Introduction

Product awareness and perception of product quality can have large effects on consumption patterns; as a result, manufacturers and marketers have developed a number of methods to both increase product awareness and to broadcast product quality to potential consumers. The methods employed to inform consumers about product quality are particularly important for experience goods, such as wine, restaurants, movies, and books, since consumers often only fully determine quality after purchase. Owing to the pervasiveness of experience goods within the marketplace, there exists a large and growing theoretical literature that examines the ways in which uncertainty regarding product quality affects consumer demand (see Akerlof, 1970; Nelson, 1970; Wiggins and Lane, 1983; and Wolinsky, 1995). Furthermore, a closely related empirical literature has developed that analyzes the extent to which product quality information affects consumer behavior. This literature examines the effect of a variety of information types and sources, including branding (Montgomery and Wernerfelt, 1992), mandatory product labeling (Jin and Leslie, 2003), and advertising (Ackerberg, 2001; Ackerberg, 2003).

One additional method used to convey quality information to consumers is through so-called experts. For example, Consumer Reports tests a large number of products each year and publishes product reviews; the Zagat guide gives quality ratings to restaurants in U.S. metropolitan areas; Ebert and Roeper review movies; and magazines such as Wine Spectator and Wine Enthusiast rate wine quality.

Most studies that analyze the impact of expert opinion on consumer demand for experience goods face a significant obstacle: products of high quality are likely to both receive high quality ratings from experts and to be of high quality. Thus, it is difficult to determine the extent to which consumer demand is affected by expert reviews, since to do so, the researcher must control for product quality, which is typically unobservable. To our knowledge, there exist only two studies that attempt to isolate the impact of expert reviews and product quality on consumer demand (Eliashberg and Shugan, 1997; Reinstein and Snyder, 2005). Further, even if expert reviews affect consumer demand for a particular good, it is unclear whether the demand impact is due to a consumer response to the quality signal in the review, or rather because the rating merely alerts consumers to the presence of the good. We know of only two papers that investigate the extent to which "any publicity is good publicity" (Sorensen and Rasmussen, 2004; Reinstein and Snyder, 2005), and they have partially conflicting results. Thus, there exists little conclusive evidence regarding the channels through which expert opinion may affect consumer demand for experience goods.

Our paper contributes to the empirical literature by examining the impact of expert opinion on retail wine purchases. To distinguish the effect of expert reviews from that of product quality, we utilize an experimental approach implemented at stores in a national retail grocery chain. Specifically, a sample of wines typically stocked by a retail store in Northern California were randomly chosen to be accompanied by wine scores from a proprietary wine scoring system, and wine opinion labels were then displayed for one month during the spring of 2006 on available bottles.

Since we received the wine price schedule for the treatment period from the retail store before we designed our treatment group and labeled the treated wines, we can be confident that the retailer's pricing schedule was not influenced by the selection of treatment wines. Using wine sales trends for previous years, we selected a series of control stores for the experiment based on a match of pre-treatment observable store, sales, and consumer demand characteristics. This allowed for the use of a difference-in-differences approach to test whether consumers respond to expert opinion, and to examine the extent to which any publicity is good publicity by investigating consumer response across wines of differing quality.

We find that sales of wines with expert opinion information significantly increase, by 25 percent on average. We also find that a wine's displayed score has a significant positive impact on demand for treated wines. Specifically, low-scoring wines experienced a decrease in demand, while high-scoring wines experienced an increase, indicating that consumers positively responded to high quality information signals and negatively responded to low quality information signals; this suggests that not all publicity is good publicity.

Interestingly, we also find that while demand significantly increased on average for treated wines, demand did not significantly change for untreated wines within the control store. There are several potential explanations for this finding. First, it may be that substitution by consumers towards treated wines and away from untreated wines was not one-for-one. That is, at least some consumers, when buying treated wines, also continued to buy untreated wines. Alternatively, substitution may have been one-for-one, but consumers who previously did not purchase wine due to a lack of quality information entered into the wine market when expert opinion information was provided. Finally, consumers may have substituted temporarily by stocking up on treated wines, or spatially by reducing the quantity of wine purchased at competing stores.

These results are similar to those of Reinstein and Snyder (2005) in which the authors exploit the timing of movie reviews by Siskel and Ebert to identify the impact of expert opinion. The authors find no overall effect of reviews, but show that positive reviews increased box office revenues for narrowly-released movies and dramas. (It remains unclear why demand increased for only this subset of movies and not for other films.) On the other hand, the results of Sorensen and Rasmussen (2004) differ somewhat from our experimental findings. They analyze the impact of book reviews in the New York Times and find that both positive and negative reviews increase book sales; however, positive reviews have a larger effect on book sales than negative reviews. Thus, expert opinion both alerts consumers to the existence of a particular book and informs consumers of a book's quality. Although it is not clear why their findings differ from those we present, it may be the result of fundamental differences between the market for books and wine, differences in the information content provided by expert reviews, or the manner in which consumers in the two markets utilize expert opinion.

The remainder of the paper is structured as follows. Section 2 discusses the design of the experiment and the data; Section 3 gives our empirical strategy; and Section 4 presents results. Section 5 concludes.

2 Experimental Design

To distinguish the effect of expert reviews from that of product quality, we utilize an experimental approach in which we randomly selected 150 wines typically stocked at a large national retail grocery chain. As with many other grocery stores, consumers choose from a large number of wines located on four to five levels of shelves along both sides of an aisle. Wine scores from a proprietary wine scoring system were displayed in the treatment store for four weeks during the month of April 2006. Since many of the wines stocked in the store did not receive scores from any of the wine rating agencies, the wines chosen for the experiment were chosen from the population of wines that did receive wine scores. Of the total of 1,089 wines sold in the treatment store in March 2006, 476, or 44 percent, received wine scores from one of several potential wine scoring agencies. By selecting 150 of those 476 as treatment wines, we treat 32 percent of the population of potential candidates and 14 percent of all wines within the store.

For each treated wine, we affixed a score label to the shelf price label. Each label prominently displayed the name of the proprietary scoring system and the wine's score. Wine scores can in theory range from 50 to 100, with 100 being the highest possible score. However, scores less than 70 are not commonly released by the rating agency, and within our sample, half the wines received scores between 82 and 86, and 90 percent received scores between 78 and 89.

For all wines sold by the retailer, we obtained weekly store-level sales data for the treatment store and 38 additional Northern Californian stores. The data provide a unique wine identifier, the name of the wine, the number of bottles sold, the pre-discount price, 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, average pre-discount price, average post-discount price, and whether a bottle of wine was discounted during a given month. For those wines for which wine scores exist, we then merge the proprietary wine score into the sales data. Due to differences between the retail chain's inventory database and those wines actually on the shelves at the time of the experiment in the store, a sub-sample of the wines we intended to treat were labeled in the treatment store.

We aggregate the data to the month-level because in general, the retailer only changes wine prices and assigns price discounts at the beginning of each month; as such, prices are constant during the month. As a result, although our data are provided at the week-level, there exists very little inter-month variation and our results are not qualitatively affected by using monthly data.

Descriptive statistics for treated wines, untreated wines with scores, and untreated wines without scores are provided in Table 1 for the pre-treatment month (March) and treatment month (April) for the treatment store. As the table indicates, few differences exist between treated wines and untreated wines for which scores exist. For example, the mean score for both treated and untreated wines is approximately equal to 83. Further, the pre-treatment difference between price, quantity, and percent of wines that are red is not significantly different across the treatment and control group. Table 1 also shows that few observable differences exist between treated wines and untreated wines for which scores are not available.

To rigorously examine the extent to which the treatment increased wine sales, we utilize three control store strategies for the difference-in-differences analysis. Control store selection and model specification are discussed in the following section.

3 Empirical Strategy

3.1 Control Store Selection

Before difference-in-differences analysis can be conducted, a control must be selected. The retailer classifies the treatment store as a high wine revenue store with wine sales accounting for a greater percentage of total sales, a greater amount of wine shelf space, more expensive wines in stock, and customers with a greater median household income relative to the median of the candidate control stores (Table 2). Starting from the total universe of 241 stores in the study region, we first restrict the analysis to stores that the retailer classifies as high revenue wine stores. This restriction reduces the number of potential control stores from 241 to 38. Given that there exists a high degree of correlation between wine sales, wine selection, and the demographics of areas surrounding such stores, this restriction ensures that the treatment stores are likely to be matched to stores with similar objective characteristics.

To further reduce the number of control stores in the control store set, we use a methodology that aims to ensure that the effect of price, discounts, and wine type on sales of wine, and pre-treatment time trends in the total number of bottles sold during each month are similar for the treatment and control stores. The latter condition is similar to a robustness check used in many difference-in-difference approaches to determine whether differences exist in the pre-treatment trends; the former condition helps to ensure that differential responses to changes in price and the existence of discounts across treatment and control stores are small or nonexistent, thus decreasing the likelihood that the estimated treatment effects are biased.

First, to determine the effect of wine characteristics on demand for each pair of treatment and candidate control stores, we estimate the reduced form equation for the number of bottles sold for the 18 months preceding the treatment intervention:

$$(1) \quad \begin{aligned} Q_{it} = {} & \alpha + \beta_1(price)_{it} + \beta_2(discount)_{it} + \beta_3(red)_{it} + \beta_4(price * red)_{it} \\ & + \beta_5(price * discount)_{it} + \beta_6(red * discount)_{it} + \beta_7(month)_t \\ & + \beta_8(month{*}price)_{it} + \beta_9(month{*}discount)_{it} + \beta_{10}(month{*}red)_{it} + \epsilon_{it} \end{aligned}$$

where Q_{it} is the number of bottles sold of wine i during month t, *price* is the average price for wine i during month t, *discount* is a dummy variable equal to one if a wine was on sale for any one week during month t, *red* indicates if a given wine is a red wine, and *month* is a vector of month fixed effects.

Second, to explore the differences in the pre-treatment trends between treatment and candidate control store pairs, we regress log quantity in the treatment store against log quantity for the candidate store for the 18 months preceding the treatment intervention:

$$(2) \quad log(Q^t_{it}) = \alpha + \beta log(Q^c_{it}) + \epsilon$$

where Q^t_{it} and Q^c_{it} are the quantities of wine i sold during time t in the treatment store and candidate store c, respectively. This method closely mimics the robustness test used in many difference-in-difference approaches to determine whether differences exist in the pre-treatment trends (see Meyer, 1995).

Candidate stores are ranked by the p-value for the Chow F-test for Eq. (1) and the estimate of β for Eq. (2). The rankings have a correlation coefficient of 0.14, indicating that the methodologies capture different processes in the data. The aggregate rank of the two tests is used to define a single control store with the lowest aggregate rank; nine control stores within the lowest 25 percent of aggregate rankings are designated as Set A.[1] The complete set of 38 high wine revenue stores are designated as Set B. Store characteristics are reported in Table 2.

The following empirical analysis utilizes each set of control stores to explore the tension between the strategies of using a single closely defined control store versus a larger set of observations from a more broadly defined set of control stores. Therefore, in estimating the effect of expert option on wine demand, we balance between minimizing unobserved differences between treatment and control stores and increasing precision.

3.2 Difference-in-Differences

We utilize a difference-in-differences approach to analyze the effect of the treatment on treated wines and to determine whether expert opinion provides quality information, or simply highlights the existence of treated wines. Specifically, we first examine the effect of the treatment on the treated wines by comparing the change in the sales

[1]See Figure A-1 in the online appendix.

of treated wines from the pre-treatment to treatment month in the treatment store to that in the control store. We do so by running the following difference-in-differences specification for the pre-treatment and treatment months on only those wines that received an expert opinion label in the treatment store:

$$(3) \quad Q_{ist} = \alpha_s + \beta_0 + \beta_1 T_{is} + \beta_2 t_{it} + \beta_3 T_{is} * t_{it} + \epsilon_{ist}$$

where Q_{ist} is the number of bottles of wine i sold in store s in time t, α_s denotes store fixed effects to control for store-specific constant factors, T_{is} is an indicator variable that is equal to one for treated wines in the treatment store and equal to zero for treated wines in the control store, and t_{it} is a month dummy that is equal to one during the treatment month and equal to zero during the pre-treatment month. We refer to this specification as Specification One. The coefficients on T_{is} can be interpreted as a treatment group specific effect, on t_{it} as a time trend common to the control and treatment stores, and the coefficient for $T_{is} * t_{it}$ can be interpreted as the true effect of the treatment.

Although useful for examining the average treatment effect of being labeled on the treated wines, Specification One does not address the extent to which the expert opinion effect is related to quality information provision versus general publicity. To examine the manner in which consumers use expert opinion information, we include interactions between the continuous score variable, the wine's price, and the treatment. If expert opinion solely provides quality information to consumers, then only those treated wines that received higher scores should experience an increase in quantity sold. Alternatively, if the only effect of expert opinion labels is to alert consumers to the existence of a wine, then the treatment should have the same impact irrespective of a wine's score.

Specification One also fails to control for potentially important covariates that, if omitted, could lead to a biased estimate of the treatment effect. For example, there exist many different types of wine and consumer demand may differ across wine varietals. To reduce the likelihood that the estimated treatment effects are biased, we include dummy variables for red wines and discounted wines, as well as price, score, and price-score-treatment interaction variables. When comparing the point estimates with and without controls, we find no significant differences, which reassures us that our experimental design is valid based on observable characteristics in the treatment and in the control stores.

Specification Two, provided below, incorporates all of the above critiques of Specification One:

$$Q_{ist} = \alpha_s + \beta_0 + \beta_1 T_{is} + \beta_2 t_{it} + \beta_3 T_{is} * t_{it} + \delta X_{it} + \gamma(X_{it} * T_{is}) + \lambda(X_{it} * t_{it}) + \theta(X_{it} * T_{is} * t_{it}) + \pi(price_{it} * score_i * T_{is} * t_{it}) + \zeta W_{it} + \epsilon_{ist} \quad (4)$$

where X_{it} is a matrix that contains the variables *price* and *score*; *price* is coded as a continuous variable that is equal to the average sale price of the wine and *score* is a value of the score that a wine received from the proprietary wine rating system. W_{it} is a matrix that contains the variables *discount* and *red*; *discount* is a dummy variable equal to one if a wine was on sale for any one week during a month and *red* indicates if a given wine is a red wine.

The primary coefficients of interest are those coefficients in the vector θ and the estimated coefficient for π. The parameters in θ allow us to examine to what extent wine characteristics, such as score and price, interact with the treatment. For example, if the coefficient of the overall treatment effect is not significantly different from zero and the estimate of the interaction between *score*, T_{is}, and t_{it} is positive and significant, then we can conclude that the treatment only increased the sales of high-scoring wines. Such a finding would support the hypothesis that expert opinion provides quality information and does not simply serve an attention-grabbing role. However, the effect of the variable *score* may differ across treated wines. If consumers who purchase expensive wines are sufficiently informed regarding wine quality, while consumers who purchase less expensive wines lack knowledge of wine quality, then quality information provision should only affect the demand for less expensive wines; the parameter π allows for such a possibility by permitting the impact of score to vary across wine price.

Although we are primarily interested in estimating the average treatment effect on the treated wines, we also estimate the average treatment effect on wines that received no label (untreated). *A priori*, it is not clear whether this estimate should be less than, greater than, or equal to zero. For example, as consumers purchase wines with expert opinion labels, they may substitute away from unlabeled wines. From a retail grocer's perspective, this is important as the pattern and extent of substitution may affect wine revenues. Alternatively, consumers who previously did not purchase wine due to a lack of information may be induced by readily available information to enter the market. While these consumers are initially drawn to labeled wines, more time in the aisle may result in increased purchases in general. To determine the extent to which the treatment affected sales of unlabeled (untreated) wines in the treated store relative to the same wines in the control store, we estimate Specification Two (Eq. 4) for the subset of untreated wines with a score.

4 Results

4.1 Differences in Means

Descriptive statistics for treated wines in the treatment store and the Set B control stores are reported in Table 3. The score row reports average scores, the quantity rows report the average number of bottles sold in a given month, the price rows report the average price in a given month, the percent discounted rows report the percentage of wines that had price discounts in a given month, and finally the percent red row reports the percentage of wines that were red wines.

The paper utilizes an unbalanced data set of wines sales data over two periods for one treatment store and multiple control stores. This is seen in the number of observations reported in the last row of Table 3; as it shows, there are 93 unique wines labeled in the treatment store and 92 of these wines are also found in the Set B control stores.[2]

The first column of Table 3 indicates that the average number of bottles of treated wines sold increased by 1.5 bottles from March to April in the treatment store, while in the second column, we see that the average number of bottles sold decreased by 1.1 bottles from March to April in the set of 38 control stores (Set B). The difference-in-differences in the means suggests that the treatment increased consumer demand for treated wines by roughly 2.6 bottles, or 27 percent.

In the remaining columns of Table 3 we investigate whether the treatment effect is different for wines with higher and lower scores. Columns 3 and 4 report summary statistics in the treatment store for treated wines, distinguishing between wines with scores lower than and greater than or equal to 81; columns 5 and 6 report the same descriptive statistics for treated wines in the control stores.[3] Table 3 illustrates that

[2] Several of the control stores used in the pooled specifications do not have observations for all wines. If the data were balanced and we used one treatment and one control, we would have 93*2+92*2= 370.

[3] We selected the cutoff of 81, as this is the level of scores for which the treatment effect is estimated to be zero. See Section 4.3 and Table 5.

in the treated store the increase in the average sales of treated wines is driven by increased demand for high-scoring wines, which increases from 9.4 to 11.5 bottles on average. Sales of low-scoring wines, however, actually decrease from 9.9 to 8.2 bottles on average. In the control stores, low-scoring wines decreased from 13.8 to 12.7, and high-scoring wines decreased from 10.8 to 9.6. In terms of the difference-in-differences, lower scoring wines had a treatment effect of a decrease of 0.7 bottles, whereas higher scoring wines experienced an average increase of 3.2 bottles.

4.2 Average Treatment Effect Regression Analysis

The examination of the average treatment effect from the first two columns of Tables 3 is supported by the regression results from Specification One (Eq. 3) provided in Table 4. Reported models utilize three alternate controls: Single Store, Set A, and Set B, which are reported in columns (1-3), (4-6), and (7-9) respectively. Continuous dependent and independent variables are transformed by the natural logarithm.

Columns (1), (4), and (7) report on the model with a constant term, store fixed effects, the treatment store dummy variable, the treatment period dummy variable, and the interaction term for the treatment store and treatment period dummy variables; the interaction term is the treatment effect and its coefficient, β_3(Eq. 3), is the primary coefficient of interest for Specification One. Other models include additional controls. Columns (2), (5), and (8) report on a model incorporating a red wine dummy and a discount dummy variable; columns (3), (6), and (9) include wine score and wine price.

The point estimates of the average treatment effect, β_3, are stable within control store sets; further, the estimates are not affected by including controls for red wine, discounts, score, and price. The average treatment effect estimate is positive and significant for the Single Store and Set B controls, with point estimates of roughly 0.40 and 0.25 respectively. These estimates correspond to approximately a 40 percent and 25 percent increase in bottles sold in the treatment store relative to the Single Store and Set B controls. This average treatment effect is roughly one third to one half of the estimated price discount effect, depending on the inclusion of price and score controls.

We obtain positive statistically significant estimates of the treatment effect when using the Single Store (columns 1-3) and Set B (columns 6-9) controls; the estimates with respect to control Set A are positive, but are not statistically significant. The

Single Store results utilize a control with the closest match as measured by the average ranking of pre-treatment demand characteristics (Eq. 1) and sales trends (Eq. 2). On the other side of the spectrum, Set B utilizes 38 control stores after making an initial first cut to exclude stores that did not generate high wine revenues. However, as the number observations increase with the size of the control set, the coefficient of determination increases, thereby indicating that the proportion of variation explained by the model increases. To be conservative, we select Set B as the preferred control as it is less likely to be influenced by any unobserved changes in any one control store and it results in an estimated treatment effect that is smaller in both magnitude and p-value.

4.3 Investigating Heterogeneity of Effects

To test whether the treatment effect is consistent across wines of different prices and scores, we include interactions between the continuous score variable, the wine's price, and the treatment. Column 1 of Table 5 provides results for Specification Two (Eq. 4) for treated wines in the treatment store, while columns 2 through 4 report on robustness checks, as well as the untreated wine regressions. All regressions reported in Table 5 utilize control Set B, the natural log transformation of all continuous variables, the full set of independent variables, and store fixed effects; all standard errors are clustered by wine.

Results indicate that wines with higher scores have larger increases in demand due to the treatment, as can be seen in the positive and significant point estimate of $Score * Tr.Store * Tr.Month$, equal to 18.3 at the five percent significance level. Using the results reported in Table 5 and the mean value of the natural log of price for wines in the treatment store, we calculate the average score at which the estimated treatment effect is zero to be approximately equal to 80. Hence, wines with scores greater or equal to 81 are estimated to experience a positive increase in demand, while wines with scores less than 81 experience a decrease in demand. These results are supported by the results of the difference in means (Section 4.1).

Additionally, we find evidence that the demand for wines in the treatment store during the treatment period is less price sensitive relative to the control, although there is a treatment effect for low-priced, high-scoring wines. Table 5 reports estimates for $Price*Tr.Store*Tr.Period$, equal to 26.7, and $Score*Price*Tr.Store*Tr.Period$, equal to -6.0; both are significant at the 10 percent level.

We draw two conclusions from these data. First, the value of a wine's displayed score has a significant positive impact on demand for treated wines; specifically, low-scoring wines experienced a decrease in demand, while high-scoring wines experienced an increase. Thus, consumers responded positively to high quality information signals and responded negatively to low quality information signals, suggesting that not all publicity is good publicity. Second, our results can also be used to investigate whether quality signals impact the price sensitivity of consumers. Results indicate that consumers exposed to quality signals are less price-sensitive relative to the consumers in the control. Additionally, there is evidence of the trade-off between price and quality for those consumers in the treatment group.

4.4 Robustness Checks

To determine whether our findings were affected by unobserved time or store effects, we performed two robustness checks, reported in columns 2 and 3 of Table 5. Column 2 reports on a Time Placebo regression; this regression is estimated with Specification Two using data from the treatment store and control store Set B for March and April of 2005, the year prior to the treatment. Results show that there is no significant effect of the false time treatment, supporting the supposition that the primary result in Column 1 is not an artifact of seasonal or other advertising trends not observed in the data. Although not reported, we assigned false treatments to all months between March 2005 and March 2006. In every case, the average treatment effect of the treatment interacted with the score is significantly less than zero.[4]

Columns 3 reports on a Store Placebo regression; this regression is estimated using the Single Store control as a placebo for the treatment store and using the remaining 37 Set B stores as controls for the March and April 2006 time period. The Single Store control was selected because it displays the demand and sales history most similar to the treatment store, thereby increasing the likelihood that a false treatment effect would be found. Regression results indicate that no treatment coefficients are significant. Additionally, we estimated Specifications One and Two using each individual Set B store as the placebo store. For Specification One, all 38 average treatment effect estimates were significantly less than zero (Figure A-2).[5] For Specification Two, the $score * treatment$ estimates were not significantly different from zero for 36 out of the 38 regressions; this finding is consistent with the sampling distribution of the test statistic under the null hypothesis for a test that reflects a

[4]In one case, there was a negative and significant $score * treatment$ estimate.

[5]In one case, there was a negative and significant estimate.

five percent probability of Type I error (Figure A-3).

4.5 Treatment Effects on the Untreated

We examine the impact of the treatment on wines in the treatment store that were not labeled using Specification Two. Column 4 of Table 5 reports that the treatment did not have a significant impact on untreated wines in the treatment store; these results suggest that consumer demand for untreated wines remained stable during the treatment period in the treatment store. There are several potential explanations for this last finding. First, it may be that possible substitution between treated and untreated wines was not one-for-one; that is average purchases of unlabeled wines remained constant while purchases for labeled wines increased. Alternatively, consumers who previously did not purchase wine due to a lack of quality information may have entered into the wine market when expert opinion was provided. Finally, consumers may have substituted temporarily by stocking up on treated wines, or spatially by reducing the quantity of wine purchased at competing stores.[6]

5 Conclusions

Unlike most previous work that examines the impact of expert opinion on consumer demand, we are able to disentangle the endogenous relationship between product quality and expert opinion provision through the use of a field experiment in a retail grocery chain. By randomly selecting wines to display expert opinion information, and through the careful selection of control stores, we are able to examine both the effect of expert opinion on the overall demand for wine and the role of expert opinion labels in providing quality information to the consumer.

During the treatment period, we affixed quality information in the form of wine score labels to the shelf price for a set of randomly chosen treatment wines. We are able to test whether the provision of information impacts the sales of treated wines relative to untreated wines, and whether the size and direction of the effect is correlated with the score received by the wine. The first test can provide evidence as to the average effect of the treatment on demand; the second test can shed light on the possibility

[6]Substitution in this context may be loosely defined as we do not formally model behavioral consumer choices. Without having more detailed individual consumer-level purchase data, we cannot test amongst the above explanations.

of an asymmetric demand response due to below- and above-average reviews. If a differentiated effect is found, it can be taken as evidence that consumers differentiate between good and bad reviews, and incorporate the quality signal in their purchases. However, if there is an average positive treatment effect, but no differential effect, it suggests that any publicity is good publicity and that reviews serve only to highlight a wine's existence.

We find that sales of wines with expert opinion information increased on average by 25 percent and that high-scoring wines experienced an increase in demand relative to low-scoring wines. These effects are only found in the treatment period, and only for treated wines. Finally, no effect was found using the placebo store that most closely matched the treatment store; although a false treatment effect was found in two out of 38 control stores, we conclude that this is consistent with a probability of Type I error at the five percent level of significance. Examining consumer behavior along a longer post-treatment period is an avenue of future research, as the treatment period of one month may not be sufficient to observe the full effect of expert opinion provision. Further, since the reviews may be read by wine buyers outside our treatment, the estimates may provide us with a lower bound of this effect.

In terms of external validity, our findings are to be taken within the context of the treatment store and its wealthier shoppers. To the extent that consumers in wealthier areas and those buying more expensive wines are likely to be more fully informed regarding wine quality than consumers in other areas, the treatment store selection should reduce the likelihood of finding a significant treatment effect. However, if wealthier shoppers care more about quality and reviews, we would have a larger likelihood of finding an effect.

Our findings suggest that expert opinion can provide quality information to consumers, as at least some consumers use such information when making purchasing decisions; rating agencies for wine and for other products such as electronics, cars, and restaurants thus likely affect consumer decisions through the provision of quality information. To the extent that certain consumers previously did not participate in the market due to a lack of product information, provision of such information may lead to market expansion as new consumers enter. Further, as quality information is distributed and consumers learn which producers are associated with high quality products, low quality producers may increase their product quality to more effectively compete with high quality producers. Both the relationship between information provision and consumer entry, and the relationship between quality information and the quality provided by producers remain interesting avenues for further research.

References

[1] Ackerberg, D. 2001. "Empirically Distinguishing Informative and Prestige Effects of Advertising," *The RAND Journal of Economics*, vol. 32, 316-333.

[2] Ackerberg, D. 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination," *International Economic Review*, vol. 44, 1007-1040.

[3] Akerlof, G. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, vol. 84, 488-500.

[4] Eliashberg, J. and S. Shugan. 1997. "Film Critics: Influencers or Predictors?," *Journal of Marketing*, vol. 61, 68-78.

[5] Jin, G. and P. Leslie. 2003. "The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Cards," *Quarterly Journal of Economics*, vol. 118, 409-451.

[6] Meyer, B. 1995. "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, vol. 33, 151-161.

[7] Montgomery, C. and B. Wernerfelt. 1992. "Risk Reduction and Umbrella Branding," *Journal of Business*, vol. 65, 31-50.

[8] Nelson, P. 1970. "Information and Consumer Behavior," *Journal of Political Economy*, vol. 78, 311-329.

[9] Reinstein, D. and C. Snyder. 2005. "The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics," *Journal of Industrial Economics*, vol. 53, 27-51.

[10] Sorensen, A. and S. Rasmussen. 2004. "Is Any Publicity Good Publicity? A Note on the Impact of Book Reviews," Working Paper.

[11] Wiggins, S. and W. Lane. 1983. "Quality Uncertainty, Search, and Advertising," *American Economic Review*, vol. 73, 881-894.

[12] Wolinsky, A. 1995. "Competition in Markets for Credence Goods," *Journal of Institutional and Theoretical Economics*, vol. 151, 117-131.

Table 1: Descriptive Statistics for Treated and Untreated Wines in the Treatment Store[7]

	Treated Wines	Untreated Wines with Scores	Untreated Wines without Scores
Score	83.83	83.81	-
	(3.29)	(3.31)	
Quantity (March)	9.51	14.93	7.55
	(10.13)	(25.94)	(13.49)
Quantity (April)	11.00	15.67	7.71
	(11.79)	(28.01)	(14.90)
Price (March)	7.63	7.71	8.92
	(4.56)	(4.87)	(6.34)
Price (April)	7.78	7.74	8.66
	(4.95)	(5.38)	(6.09)
% Discounted (March)	92.68	97.83	73.60
% Discounted (April)	92.77	97.36	74.31
% Red	64.36	54.79	60.60
# of Observations	93	292	522

[7]For all continuous variables, we report the mean and standard deviation. Quantity is the average number of bottles, Price is the average price, Percent Discounted is the percentage of wines that were discounted, and Percent Red is the percentage of wines that were red wines.

Table 2: Store Characteristics[8]

	Treatment Store	Single Control Store	Control Stores (Set A)	Control Stores (Set B)	All Possible Control Stores
Wine Sales Rank	36	31	31	27.5	136
Wine Sales (2005 $)	604,863	639,459	647,202	693,352	317,847
% Wine Sales of Total Grocery Sales	9.0	10.3	7.2	6.2	4.5
% Sales Premium Wine	9.7	10.9	8.2	8.5	4.4
Median Household Income in 2005	140,618	129,274	106,692	99,695	68,623
Shelf Space Linear Feet	510	285	390	38	375
# of Stores	1	1	9	33	267

[8]Store characteristics for all stores are reported for the 24-week period ending on 1/15/2006. Percent Sales Premium Wine is the percent of sales that were obtained from the sale of bottles priced greater than $8. Sales data are provided by Infoscan and median household income data are provided by the retailer.

Table 3: Descriptive Statistics for Treated Wines in the Treatment and Control Stores[a]

	Treated Store (1)	Control Set B (2)	Treated Store Score<81 (3)	Treated Store Score>=81 (4)	Control Set B Score<81 (5)	Control Set B Score>=81 (6)
Quantity (March)	9.51	11.25	9.92	9.44	13.76	10.79
	(10.13)	(14.22)	(7.43)	(10.56)	(12.98)	(14.39)
Quantity (April)	11.00	10.14	8.17	11.48	12.70	9.61
	(11.79)	(14.58)	(8.29)	(12.27)	(12.98)	(14.85)
Price (March)	7.63	8.27	4.37	8.18	9.12	8.97
	(4.56)	(5.26)	(1.70)	(4.67)	(5.47)	(5.36)
Price (April)	7.78	8.92	4.90	8.27	5.22	9.70
	(4.95)	(5.62)	(1.45)	(5.16)	(1.66)	(5.84)
Score	83.76	84.00	78.38	84.68	78.53	85.06
	(3.26)	(3.52)	(2.20)	(2.22)	(2.10)	(2.65)
% Red	64.85	64.36	83.33	61.70	78.88	61.55
# of Unique Wines	93	92	13	80	13	79

[a]The mean and standard deviation are provided for all continuous variables. Quantity is the average number of bottles sold, Price is the average price, and Percent Red gives the percentage of wines that were red wines. Column 1 provides summary statistics for treated wines in the treatment store; column 2 provides summary statistics for the 38 stores in Control Stores Set B. Columns 3 and 4 report summary statistics for the treatment store treated wines for wines with scores lower and higher than 81. Columns 5 and 6 report summary statistics for control stores treated wines by scores. Eighty-one is defined as the cut-off, as this is the level for which the treatment effect is estimated to be zero in Table 5.

Table 4: OLS Results for Specification One[a]

Control Store		Single Store			Set A			Set B	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Tr. Store	-0.45***	-0.46***	-0.47***	-0.06	-0.10	-0.11	-0.11	-0.15	0.51***
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.14)
Tr. Period	-0.26	-0.24	-0.26	-0.04	-0.04	-0.01	-0.12	-0.11	-0.08
	(0.19)	(0.18)	(0.18)	(0.11)	(0.11)	(0.09)	(0.10)	(0.10)	(0.09)
Tr. Store*Tr. Period	0.41*	0.39*	0.41**	0.19	0.18	0.17	0.27**	0.26**	0.23**
	(0.21)	(0.20)	(0.20)	(0.13)	(0.13)	(0.12)	(0.12)	(0.12)	(0.11)
Red Wine		-0.16	-0.14		-0.13	-0.06		-0.08	-0.01
		(0.19)	(0.20)		(0.20)	(0.19)		(0.19)	(0.18)
Discount		1.09***	0.97***		0.91***	0.66***		0.78***	0.51**
		(0.26)	(0.29)		(0.19)	(0.21)		(0.21)	(0.21)
Score			-1.76			-2.09			-2.77
			(2.45)			(2.41)			(2.10)
Price			-0.16			-0.37**			-0.41**
			(0.19)			(0.17)			(0.16)
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Store Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.02	0.11	0.13	0.04	0.11	0.17	0.08	0.13	0.21
# of Observations	229	229	229	807	807	807	2,909	2,909	2,909
# of Control Stores	1	1	1	9	9	9	38	38	38
# of Unique Wines	93	93	93	93	93	93	93	93	93

[a]Results are for Specification One for treated wines in the treatment and control store for the pre-treatment and treatment month. Standard errors are clustered by 93 wine ids and are given in brackets. *, **, and *** indicate that a point estimate is significant at the 10, 5, and 1 percent levels respectively. Continuous dependent and independent variables are in natural logs. We use three alternative sets of controls: a Single Store, a Set A of 9 stores, and a Set B of 38 stores. Not all wines labeled in the treatment store were stocked at all control stores during the treatment period.

Table 5: OLS Results for Specification Two[9]

	Treated Wines Primary Result (1)	Treated Wines Time Placebo (2)	Treated Wines Store Placebo (3)	Untreated Wines (4)
Tr. Store*Tr. Period	-80.89**	-21.59	-16.44	-32.77
	(37.06)	(66.42)	(88.96)	(29.17)
Score	2.04	5.32	-2.54	-0.24
	(6.76)	(15.23)	(6.52)	(4.20)
Price	9.96	-1.08	3.37	0.18
	(11.97)	(27.11)	(11.29)	(8.15)
Score*Tr. Store*Tr. Period	18.32**	4.91	-3.46	7.48
	(8.38)	(15.03)	(20.05)	(6.58)
Price*Tr. Store*Tr. Period	26.73*	0.68	-19.74	14.73
	(15.75)	(28.20)	(49.28)	(13.04)
Score*Price*Tr. Store*Tr. Period	-6.03*	-0.19	4.29	-3.35
	(3.55)	(6.36)	(11.08)	(2.94)
Red Wine	0.01	-0.30	-0.02	-0.31***
	(0.18)	(0.23)	(0.16)	(0.10)
Discount	0.53**	0.64***	0.51**	0.40
	(0.22)	(0.22)	(0.21)	(0.30)
Store Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.22	0.35	0.21	0.17
# of Observations	2,909	1,549	2,448	8,924
# of Control Stores	38	38	37	38
# of Unique Wines	93	79	70	261

[9] All results are for Specification Two. Column 1 reports results for the treatment store and Control Set B during the treatment and preceding control period. Column 2 reports Time Placebo results for data from March and April, 2005. Column 3 reports Store Placebo results for a "placebo-treatment store" designated as Control Store 1. Column 4 reports results for untreated wines in the treatment store and Control Set B during the treatment and preceding control period. Continuous dependent and independent variables are in natural logs. Estimation results for additional treatment store and period interaction terms are suppressed. Standard errors are clustered by wine and are given in brackets. *, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.