The Effects of Expert Opinion on Consumer Demand for Goods with Credence Attributes: Evidence from a Natural Experiment

Chen Zhen

RTI International

Xiaoyong Zheng

North Carolina State University

January 2015

Abstract

Little research has been dedicated to examine the effect of expert opinion on consumer demand for goods with credence attributes. Taking advantage of a natural experiment, we use difference-in-differences approach to estimate the effect of the NuVal score, an expert opinion of the healthfulness of the food product, on consumer demand. Our results show that posting the NuVal score increases yogurt sales and the effect is larger for products with higher scores. This implies that experts are opinion influencers. Consumers respond positively to the better-nutrition signals provided by the NuVal score. The publicity effect of the new label is also at work.

Keywords: Expert Opinion; Credence Goods; NuVal

JEL Classification Codes: D18; L15; M31

1 Introduction

Economists have long been interested in identifying and measuring the effects of various determinants of consumer demand. Though a vast literature has attempted to estimate the effects of price, advertising (Ippolito and Mathios 1990; Ackerberg 2001), seller reputation (Cabral and Hortacsu 2010; Elfenbein, Fisman and McManus 2012), quality disclosure (Mathios 2000; Dranove and Jin 2010) and other factors, less attention has been paid to the role of expert opinion. In markets where the information asymmetry problem is severe, assurance from independent sources such as experts, third-party certifiers and other consumers may help alleviate the problem and improve market efficiency.

In this study, we investigate the effect of an expert opinion on the healthfulness of the product on consumer demand. The nutritional label, NuVal, scores food products on a scale from 1 to 100 based on the Overall Nutritional Quality Index (ONQI) algorithm. The ONQI algorithm was developed by a multidisciplinary team of public health and nutrition scientists independent of food industry interests and was also recently validated by independent researchers (Katz et al. 2010).[1] The NuVal nutritional scoring system takes the health effects of more than 30 nutrients into account and aims to rank foods by their relative healthfulness (Katz et al. 2007). Nutrients with generally favorable effects on health such as fiber, vitamins and omega-3 fatty acids are placed in the numerator, where higher values increase the NuVal score. Nutrients with generally unfavorable effects on health such as saturated fat, trans fat, sodium, sugar and cholesterol are placed in the denominator, where higher values decrease the NuVal score. In addition to nutrients, the ONQI algorithm also takes into account other key nutrition factors that measure the quality and density of nutrients, as well as the strength of their association with specific health conditions. In sum, the NuVal score represents a validated expert opinion on the healthfulness, a credence attribute, of the food product. Products with higher NuVal scores are considered by the ONQI expert panel to be more nutritious than products receiving lower scores.

[1] Chiuve et al. (2011) used ONQI to score the diet quality of over 100,000 men and women who started as healthy individuals in two longitudinal surveys spanning more than 20 years. The authors found that baseline diets that were scored lower by the ONQI algorithm are significantly associated with higher risks of chronic disease later in the surveys.

Taking advantage of the natural experiment where a grocery retailer adopted a new price tag with the NuVal score, we compare changes in the sales of yogurt products with NuVal scores before and after the adoption in the adopting store to those of the same products in stores in the same market that did not adopt the NuVal scoring system. This difference-in-differences approach allows us to obtain a clean estimate on the effect of the NuVal score, an expert opinion of the healthfulness of the product, on consumer demand. For any food product, consumers may have already possessed some knowledge about the healthfulness of the product through tasting and reading information printed in the Nutrition Facts Panel (NFP).[2] The effect from this part of the knowledge is captured by the product fixed effects in our regression. Our analysis reveals the effect from the additional knowledge about the healthfulness of the product that cannot be learned easily by the consumers themselves and conveyed by the NuVal score.

We find that posting the NuVal score increases yogurt sales and the effect is larger for products with higher scores. This implies that experts are opinion influencers. Consumers respond positively to the better-nutrition signals provided by the NuVal score. The publicity effect of the new label is also at work. If there were no publicity effect, then demand for the lower-scoring yogurt products should decrease as their NuVal scores indicate these products have lower nutritional value. The NuVal score attracts consumers' attention to these products and may also lead consumers to perceive the lower-scoring products as having better nutritional value than those products without a NuVal score.

This study is not the first to estimate the effect of expert opinion on consumer demand. Eliashberg and Shugan (1997) find reviews from movie critics are correlated with late and cumulative box revenue but not with early box revenue, suggesting there is no evidence that experts are opinion influencers. Reinstein and Snyder (2005) use difference-in-differences method to study the causal effect of the reviews from two leading movie critics. They find positive reviews have positive effect on dramas and narrowly-released movies, but not movies in general. Berger, Sorensen and Rasmussen (2010) use both empirical as well as experimental approaches to examine the effect of book reviews. They find the interesting result that in some

[2] For example, an average consumer probably knows that a product might not be very healthy if it tastes very sweet or salty.

situations, even negative reviews can have a positive effect. Hilger, Rafert and Villas-Boas (2011) conduct a field experiment to study the effect of wine scores from a proprietary scoring system in the supermarket retail setting. They find a positive and statistically significant effect and that expert opinion transmits quality information as opposed to solely shelf visibility. Our paper differs from these studies in that we focus on a credence attribute (i.e., nutritional quality) of an experience good, while previous studies focus on the effect of expert opinion on consumer demand for experience goods without credence attributes. Experience and credence goods or attributes (Nelson 1970; Darby and Karni 1973) share the common feature that quality is unknown to consumers before purchase and hence expert opinions help alleviate the information asymmetry problem. They differ in that after purchase, consumers learn the quality of experience goods without credence attributes, but still do not know the quality of credence goods or credence attributes of experience goods. Ex ante, it is not clear whether expert opinions should have a larger or smaller effect on consumer demand for goods with credence attributes compared with that of experience goods without credence attributes. On one hand, there is no way or at least easy way for consumers to learn the quality of the credence attribute by themselves even after purchase and hence they might be more likely to rely on the recommendations by experts. On the other hand, there is also no easy way for the experts to prove to consumers that their opinions are correct both before and after purchase and hence consumers may not believe in the experts and follow their suggestions. Therefore, it is worthwhile to investigate the effect of expert opinion on consumer demand for goods with credence attributes, especially given the fact that the evidences reviewed above for goods without credence attributes are mixed.

This paper also builds on the large literature on food labeling.[3] Mathios (2000) examines the effect of the Nutrition Labeling and Education Act (NLEA) of 1990, which mandated the display of the Nutrition Facts Panel (NFP), on consumer demand for salad dressings. He finds products with low nutritional quality experienced declines in market shares after the mandatory disclosure of nutrient content and the effect is heterogeneous across consumers. Kiesel and Villas-Boas (2013) use a field experiment to examine the effects of several alternative shelf labels on consumer demand for microwave popcorn. These labels aim to reduce consumer information

[3] For brevity purpose, we only discuss the most relevant ones here.

cost by highlighting and summarizing information from the NFP (e.g. low fat; no trans fat; no calorie). They find labels of low calories and no trans fat increase sales, while the label of low fat decreases sales. Furthermore, combining multiple claims into one label increases consumer information cost and does not affect sales significantly, compared with the status quo. Zhu and Huang (2014) use difference-in-differences method to examine the effect of the Facts Up Front (FUF) nutrition labels on consumer purchase of ready-to-eat cereals. The FUF labels do not assess the nutritional quality of the products. Instead, these labels present levels of selected nutrients on the front of the package. Because the same information is available from the NFP on the back or side of the package, FUF labels are similar to those used by Kiesel and Villas-Boas (2013) in that they aim to reduce information acquisition costs but do not provide expert opinion on the nutritional quality of the products. They find that FUF labels induce consumers to consume less sodium and fewer calories and the effect is different for households with different levels of education. We contribute to this literature by studying a different kind of label, a label that conveys expert's opinion on the healthfulness of the product.

The rest of the paper is organized as follows. The next Section provides some background information on the motivation and development of the NuVal scoring system. Section 3 describes the data. Our empirical strategy is detailed in Section 4. Section 5 presents and discusses the results and the final Section concludes.

2 Information Disclosure and Nutrition Labels

Information disclosure policies have played a major role in U.S. nutrition policies aimed to reduce obesity, prevent diet-related noncommunicable diseases, and promote healthy eating. Prominent examples include the NLEA of 1990 mandating standardized NFP on most packaged foods by 1994 and the required disclosure of trans fat content on NFP by 2006. These labeling regulations would be most effective if people correctly process the disclosed nutrition facts, which are often in lengthy format on the back or side of the package, and act on this information by choosing healthier products. However, for an average person who makes over 200 daily food decisions (Wansink and Sobal 2007), it may be challenging to review and process all of this labeling information. Indeed, the literature has documented that food label use varies significantly across sociodemographic subgroups (Ollberding et al. 2010) and that diet and health knowledge is one of the strongest predictors of label use (Drichoutis et al. 2006). In addition, over the decade following NLEA's full implementation, consumer use of most nutrition labels had declined (Todd and Variyam 2008).

The obesity epidemic that has escalated post-NLEA and other health concerns associated with food choices motivated the search for more effective labeling strategies that supplement the NFP. In response, the food and beverage industry, through its Grocery Manufacturers Association and Food Marketing Institute, introduced the FUF labels.[4] Food and beverage manufacturers voluntarily decide whether or not to adopt FUF labels. A basic FUF label lists calories per serving and information about saturated fat, sodium and sugar – nutrients the Dietary Guidelines for Americans recommend limiting, on the front of the food package. In addition, manufacturers may also include information on two nutrients to encourage. These nutrients – potassium, fiber, protein, vitamin A, vitamin C, vitamin D, calcium and iron – are under-consumed and are needed to build a "nutrient-dense" diet, according to the Dietary Guidelines for Americans. It is worth noting that the FUF labels do not provide consumers with new information that is already in the NFP. Instead, they simply highlight information on a few key nutrients. Hersey et al. (2013) reviewed the literature on FUF labels. Studies in the literature have found that FUF labels reduce consumer's time to process nutrition information, but results regarding whether FUF labels can help consumers choose the healthier products are mixed.

Around the same time, private enterprises heeded the business opportunities with providing expert assessment of product nutritional quality at point of purchase. NuVal, licensed through NuVal LLC, is such a label. The NuVal label is essentially a new price tag with the NuVal score for the product on it. See Figure 1 for examples of the NuVal labels. Its underlying scoring system, ONQI, was developed during the two-year period from 2005 to 2007 by a panel of public health and nutrition scientists. NuVal LLC started scoring food products using the scoring system in the middle of 2008. In the same year, it formed a partnership with Topco Associates LLC, which is a private business consulting company jointly owned by many grocery retail chains, to market the NuVal labels to grocery stores. In January 2009, Price Chopper and Hy-Vee became the first two retail chains that adopted the NuVal labels. Due to the large

[4] More information is available at http://www.factsupfront.org/. Last access date: 11/26/2014.

number of food products on the market, NuVal scored food products over time. By November 2010, it had scored 75,000 food products. By November 2014, it had been adopted by thirty-three retail chains and several public school systems throughout the U.S.[5]

In addition to NuVal, Guiding Stars is another multiple-level summary nutrition shelf label system on the U.S. market (Rahkovsky et al. 2013). Similar to NuVal, Guiding Stars rates the nutritional quality of foods using a four-point system. The algorithm behind the Guiding Stars assigns scores based on the levels of nutrients recommended by key scientific bodies (e.g., the Dietary Guidelines for Americans) to limit or to encourage per 100 kcal of dietary energy (Fischer et al. 2011). Food products in the top, second, and third tertile of the positive range of the score receive three, two, and one star on the shelf tag, respectively, where three stars indicate the highest degree of healthfulness. Foods receiving zero or negative scores are unstarred and receive no shelf tag. Different from NuVal, the Guiding Stars label is a standard-alone label affixed on the shelf, not part of the price tag. By November 2014, the Guiding Stars label had only been adopted by five grocery retailers.

In 2012, the Institute of Medicine reviewed various nutrition labels currently in use including the FUFs, NuVal and Guiding Stars. It concluded with a recommendation to develop a government-sponsored summary multiple-level nutrition symbol that goes on the front of the package and provides a clear ranking of the healthfulness of the products (IOM 2012). Such a label would combine features from both the FUFs and labels such as NuVal and Guiding Stars. The IOM report encourages government regulators to shift from the current cognitive approach of providing more written information on nutrition facts to an interpretive one that provides simple, direct, and science-based guidance to consumers on the nutritional quality of the products.

3 Data

This section describes the data used in the subsequent empirical analysis. From the IRI Academic Data Set (Bronnenberg et al. 2008), we obtained the scanner data for all yogurt products sold in the universe of six grocery stores in a small town in Midwest. One of these six

[5] Most information in this paragraph was taken from NuVal's online newsroom, http://www.nuval.com/News. Last access date: 11/26/2014.

stores is owned by a regional grocery chain, which adopted the NuVal labels in August 2010. Among the rest, two are owned by a local food co-op, one by another regional grocery chain and two each by a local independent owner. None of these five stores adopted the NuVal or the Guiding Stars labels during our sample period. Therefore, the store that adopted the NuVal label is the treatment store and the other five stores are the control stores in our empirical analysis.

The scanner data is at the Universal Product Code (UPC) level. For each UPC, we have information on the weekly sales, the number of units sold, whether the price mark-down is larger than 5%, as well as advertising and promotion activities in the store. We calculated the average weekly price for each UPC by dividing the weekly sales by the total number of units sold in that week. In addition, the scanner data provide information on the following product characteristics for each UPC: the producer, whether it is organic, whether it is a Greek yogurt, whether it is soymilk based and whether it is a yogurt fluid drink. We chose to focus on the yogurt products because it is a well-defined food category with many varieties, providing enough variation in the NuVal score variable, the variable of key interest.

The NuVal scores for the UPCs are obtained from NuVal LLC, NuVal's licensing agency. Due to the large number of food items sold in grocery stores, NuVal LLC scored the products in batches over time, rather than scoring all of them at the same time. NuVal released scores for the first batch of yogurt products in January 2009. By August 1st, 2010, it had scored 918 different yogurt products and by October 25, 2013, it had scored 2,372 different yogurt products. We matched the NuVal scores with the scanner data using the unique product identifier, that is, the UPC.

Since our treatment store adopted the NuVal scores in August 2010, we define September 2010 to December 2010 as the treatment period and April 2010 to July 2010 as the control period for our empirical analysis. As yogurt manufacturers frequently introduce new products and terminate old ones, we chose a relatively small bandwidth, that is, 4 months, to minimize the effect of changes in product composition on consumer demand. During the treatment period, 367 UPCs were sold at the treatment store, while 374 were sold during the treatment period. For 155 of them, the NuVal scores had already been available by August 1st, 2010 and all of these

155 UPCs were sold during both the treatment and control periods. Out of these 155 UPCs, 130 were sold in the control stores during both the treatment and control periods. These 130 UPCs are therefore the treated UPCs in our difference-in-differences analysis. To shed light on the differences between the treated and untreated UPCs, Table 1 presents the summary statistics for the treated and other UPCs in the treatment store during the control period. As can be seen from the Table, on average the treated UPCs had a lower unit price and larger quantities sold.[7] This is consistent with the fact that NuVal LLC scored the more popular products first. Also, the treated UPCs are more likely to be produced by General Mills (owner of the brand Yoplait) and Dannon, the two leading firms in this industry, and less likely to be Greek or organic (at that time, the Greek yogurt was not as popular as of today).

Our empirical analysis below focuses on the 130 treated UPCs defined above, comparing changes in their sales quantities before and after the NuVal labels were adopted in the treatment store versus those in the control stores. From Table 1, we can see that the NuVal scores for the 130 UPCs range from 23 to 100, with an average score of 51.8. This shows that the treatment store not only posted the NuVal score for high-scoring products, but also for low-scoring products. This is because, unlike the stand-alone Guiding Stars label, NuVal scores are integrated into the price tag and will be posted for all products that NuVal has already scored, once the retail chain adopts the NuVal price tag system. Table 1 also shows that 52% of the 130 UPCs are produced by General Mills, 33% by Dannon and the rest by other small companies. In terms of product features, 4.6% of them are organic, 5.4% are Greek, 3.1% are soymilk based and 5.4% are fluid drinks. On average, about 36 units were sold every week for each UPC, with an average price of $1.43 per unit. There were very little advertising and promotion activities in the store. Only 1.22% and 0.79% of the observations (the unit of observation is one UPC in one week in one store) come with a minor display and a medium-size ad respectively and there was zero occurrence of large display or ads of other sizes.

To examine the differences between the treatment store and the control stores, Table 2 provides summary statistics for the yogurt products sold in these stores during the control period. The

[7] A lower unit price does not necessarily mean the product is cheaper as the package size differs for different products.

Table shows that the treatment store was the market leader for the yogurt products in this small town during our sample period. It carried 130 treated UPCs and 244 other UPCs during the control period. The five control stores carried 83 to 124 treated UPCs and 92 to 204 other UPCs. The treatment store also sold more units per product. For the treated UPCs, it sold about 36 units per UPC per week, while the numbers at the control stores range from 9 to 34. For other UPCs, it sold about 19 units per UPC per week, while the numbers at the control stores range from 7 to 19. Also, the unit prices charged by the treatment store for the treated UPCs were similar to those of the control stores except control store 4, which carried a significant less number of yogurt products. The average price per unit at the treatment store for the treated UPCs was $1.43, while the unit prices at the four control stores other than control store 4 range from $1.42 to $1.46. The average price per unit at the treatment store for other UPCs was slightly higher than those of the control stores at $2.12, while the unit prices at the control stores range from $1.45 to $1.89. Part of this difference comes from the differences in the number of products offered. Overall, although the treatment store was clearly the market leader in this market, the differences between the treatment store and four out of the five control stores were not large in terms of product variety and sales.

3.1 Differences in Means

Before turning to the regression analysis, we first use summary statistics to examine the treatment effect. Table 3 provides the average weekly quantity sold and price charged for the treated UPCs before and after the NuVal labels were adopted and at both the treatment store and the control stores. The first two columns of the Table report the summary statistics for all treated UPCs, regardless of the NuVal scores. As we can see, in the treated store, the average quantity sold for each treated UPC increased from 36.18 units to 45.00 units after the NuVal labels were adopted. The average unit price charged changed very little with a slight increase less than 1 cent. During the same time, in the control stores, the average quantity sold for each treated UPC decreased from 24.65 units to 19.34 units, with the average unit price decreasing for about 3 cents. The difference-in-differences in the means suggests that posting NuVal labels increased consumer demand for the treated UPCs by 14.13 units, or about 39% of 36.18, the average quantity sold in the treatment store before the NuVal labels were adopted.

In the remaining columns of Table 3, we investigate whether the treatment effect is different for UPCs with higher and lower NuVal scores. Columns 3 and 4 report summary statistics in the treatment store for the treated UPCs with scores higher than or equal to 50,while columns 5 and 6 report the same descriptive statistics for treated UPCs with scores lower than 50.[8] Table 3 illustrates that in the treated store, the increase in the average quantity sold in the treated UPCs is driven by an increased demand for high-scoring UPCs, which increased from 51.17 units to 70.40 units. Sales of the low-scoring UPCs, however, decreased from 21.91 to 20.83 units on average. In the control stores, high-scoring UPCs decreased from 31.20 to 24.54, while low-scoring UPCs decreased from 17.55 to 13.59. In terms of the difference-in-differences, higher-scoring UPCs had a treatment effect of an increase of 25.89 units, whereas lower-scoring UPCs experienced an average increase of 2.88 units.

Although the differences in means results are informative, these simple comparisons do not take into account the effects of many observed as well as unobserved factors on consumer demand. Our empirical strategy below uses regression analysis to control for these factors and analyze the treatment effect in a more rigorous way.

4 Empirical Strategy

We use difference-in-differences regressions to examine the effect of NuVal labels on the treated UPCs, controlling for many other observed and unobserved factors that affect consumer demand for yogurt products. We run the following regression using data on the treated UPCs in both the treatment and the control stores and during both the control and the treatment periods,

$$\log Q_{ist} = \beta_0 + \beta_1 D_s + \beta_2 D_t + \beta_3 D_s * D_t + \beta_4 X_{ist} + \alpha_i + \gamma_s + \lambda_t + \varepsilon_{ist} \tag{1}$$

where $\log Q_{ist}$ is the logarithm of the number of units sold in store s during week t for UPC i, D_s is a dummy variable indicating the treatment store, D_t is a dummy variable indicating the treatment period and X_{ist} is a set of price, advertising, and promotion activities variables that may or may not vary across store and time. α_i is the time- and store- invariant UPC fixed effect controlling for unobserved product characteristics that are likely to affect consumer demand.

[8] We selected the cutoff point of 50 for two reasons. First, it is close to the mean score for the treated UPCs, that is, 51.8 as reported in Table 2. Second, as the scores are designed to range from 0 to 100, 50 is a natural cutoff point as consumers may believe that UPCs with scores higher than 50 are healthy ones and those below 50 are unhealthy ones.

Among other things, it captures the part of knowledge consumers already know about the healthfulness of the yogurt product through past experience and reading information printed in the NFP. γ_s is the store fixed effect controlling for the UPC- and time-invariant store-specific factors that are likely to affect consumer demand. λ_t is the week fixed effect controlling for store- and UPC-invariant time-specific factors that are likely to affect consumer demand, and finally, ε_{ist} is the error term. The coefficient of interest is β_3, which can be interpreted as the average treatment effect on the treated. Note that in (1), the term $\beta_1 D_s$ is subsumed in the store fixed effects γ_s and the term $\beta_2 D_t$ is subsumed in the time fixed effects λ_t. Hence, coefficient estimates for β_1 and β_2 will not be reported.

Although useful for examining the average treatment effect of the NuVal label on the treated UPCs, (1) does not address the extent to which the NuVal label effect is related to nutritional information provision by the experts versus a salience effect from the label. To distinguish between the two effects, we interact NuVal score and unit price with the treatment indicators. If consumer opinion is influenced by the experts, then treated UPCs with higher scores should experience a higher increase in quantity sold compared with treated UPCs with lower scores. Alternatively, if the only effect of the NuVal label is to alert consumers to the existence of a product, i.e. the salience effect, then the treatment should have the same impact across products regardless of their NuVal scores.

More specifically, we estimate the following equation,

(2) $\log Q_{ist} = \beta_0 + \beta_1 D_s + \beta_2 D_t + \beta_3 D_s * D_t + \theta_1 S_i * D_s * D_t + \theta_2 \log p_{ist} * D_s * D_t + \theta_3 \log p_{ist} * S_i * D_s * D_t + \beta_4 X_{ist} + \alpha_i + \gamma_s + \lambda_t + \varepsilon_{ist}$,

where S_i is the NuVal score for UPC i and $\log p_{ist}$ is the logarithm of the unit price for UPC i in store s during week t. The θ parameters allow us to examine how the treatment effect varies with price and the NuVal score of the UPC. For example, if θ_1 is estimated to be positive and statistically significant, then we can conclude that the treatment effect is higher for UPCs with higher NuVal scores. Such a finding would lend support to the hypothesis that experts are opinion influencers and consumers follow their suggestions. The treatment effect may also vary across UPCs based on price. If θ_2 is estimated to be positive and statistically significant, then we can conclude that the treatment effect is larger for more expensive UPCs. Such a finding would be consistent with the hypothesis that consumers of more expensive yogurt products are more attentive to and influenced by the NuVal label than consumers of less expensive yogurts. Finally, a statistically significant θ_3 indicates that the difference in treatment effect between products with different prices identified in θ_2 is at least partly due to the expert opinion conveyed through the NuVal scores.

4.1 Identification

The validity of our empirical strategy above depends critically on two identification assumptions. First, estimation of the average treatment effect depends on the assumption that treatment is independent of potential outcomes (Imbens 2004), that is, the treatment variables in (1) and (2) are exogenous. This assumption would be violated if the treatment store timed when to introduce the NuVal labels based on some unobserved demand factors. We believe this is a highly unlikely scenario for two reasons. First, the decision to adopt the NuVal labels was a strategic decision made at the retail chain level to cater more to health conscious consumers, not by the manager of the treatment store. Second, even if the decision was made in response to changes in unobserved demand factors, the response was to changes in the overall demand at all stores owned by the chain, not to the idiosyncratic demand shocks at the treatment store.

Second, since both (1) and (2) are demand functions, supply-side variables like price and advertising are potentially endogenous. Endogeneity bias could arise if stores condition on time-varying and store-specific unobserved factors when they make their decisions on price and advertising activities. However, Dube, Hitsch and Rossi (2010) argue that price variation in scanner data is mainly across brands, which is controlled in our regressions using the UPC fixed effects. Only a small percentage of variation is explained by store and time effects. Hendel and Nevo (2013) also argue that potential correlation between prices and the error term is not a major concern in their demand estimation using the scanner data. Therefore, we follow their approaches and use OLS rather than instrumental variable approaches to estimate (1) and (2), after including UPC, store and week fixed effects in the empirical models. In our robustness checks below, we will also examine how out results change when these potential endogenous variables are omitted from the regressions.

5 Results

The final sample of dataset used to estimate equations (1) and (2) come from data on the 130 treated UPCs from the 4-month control period and the 4-month treatment period in both the treatment and the control stores. The total number of observations is 23,758. The unit of observation is one UPC in one store during one week. Definitions and the summary statistics for the variables used in estimation are displayed in Table 4. Table 4 shows that about 20% of the observations are from the treatment store and 50% of the observations are from the treatment period. As a result, about 10% of the observations are from the treatment store during the treatment period. There were very little advertising and promotion activities. Only about 2% of the observations come with a large size ad and 4% of the observations come with a medium size ad. Small size ad was never used. On the other hand, the strategy of a large price mark-down was often used, about 22% of the observations having a price mark-down larger than 5%.

Estimation results for (1) are reported in Table 5. Two specifications are estimated. In the first specification, only the price variable is included as part of the X_{ist} variable. In the second specification, variables for the advertising and promotions activities and the price reduction indicator are also included. Week, store and UPC fixed effects are included in both specifications. The standard errors are adjusted for clustering at both the UPC and the week levels. Therefore, the obtained standard errors are robust to correlations and heteroskedasticity among observations for the same UPC and/or during the same week.

As results from both specifications are similar, we focus our discussion on the first specification. First, the R^2 is close to 0.75, indicating we have included the most relevant variables influencing yogurt sales in our regressions as they together explain a large fraction of the variation in the dependent variable. Second, the estimated price elasticity is -2.40 and the estimate is significant, indicating that demand for yogurt products at the UPC level is quite elastic. Third, the estimated average treatment effect on the treated is 13.67% and statistically significant, meaning that on average, posting the NuVal label boosts sales of the treated UPCs at the treatment store by 13.67% relative to the control stores. This effect is equivalent to that of a 5.7% price reduction for treated UPCs.

5.1 Heterogeneity of the Treatment Effect

To test whether the treatment effect varies across UPCs with different NuVal scores and prices, we interact the NuVal score variable, the price variable and the interaction between the two, with the treatment store and period variables. Regression results from (2) are displayed in Table 6. Same as in Table 5, the first specification does not include advertising and promotion activity variables, while the second specification includes them. Again, as results from both specifications are quite similar, we focus our discussion below on the results from the first specification.

The estimated treatment effect is,

(3) $-0.0445 + 0.0031 * S_i + 0.2466 * \log p_{ist} - 0.0016 \log p_{ist} * S_i$.

Using the mean value of the logarithm of price for UPCs in the treatment store during the treatment period, the treatment effect becomes a function of the NuVal score only,

(4) $-0.0209 + 0.00295 * S_i$.

(4) shows that the treatment effect is larger for UPCs with a higher NuVal score. The lowest NuVal score for the treated UPCs in our dataset is 23. The estimated treatment effect for a UPC with a score of 23 is 0.0470. This means that compared with the demand in the control stores, demand for this UPC increases by 4.70% with the posting of the NuVal label. The highest NuVal score for the treated UPCs in our dataset is 100. The estimated treatment effect for a UPC with a score of 100 is 0.2714. This means that compared with the demand in the control stores, demand for this UPC increases by 27.14% with the posting of the NuVal label.

Therefore, the treatment effect is positive for all treated UPCs and the effect is larger for UPCs with higher NuVal scores. This result implies that expert opinion influences consumers, who respond positively to the better-nutrition signals provided by the NuVal label and increase their demand for higher-scoring UPCs more. The publicity effect is at work as well. If there was no publicity effect, then demand for the lower-scoring UPCs should decrease as the NuVal label tells the consumers these products have low nutritional value. The NuVal label attracts consumers' attention to the treated UPCs and may also lead consumers to (wrongly) believe that the lower-scoring products are better than those products that haven't been scored yet. As a result, demand for the lower-scoring products also increases.

Additionally, the estimated results reveal how consumer price elasticity changes with treatment and the NuVal score. The estimated price elasticity is,

(5) $-2.4129 + 0.2466 * D_s * D_t - 0.0016 * S_i * D_s * D_t$.

The result that the estimate for θ_2 is 0.2466 and significant at 1% level has two implications. First, it shows that consumers of more expensive yogurt products are more attentive to and influenced by the NuVal label than consumers of less expensive yogurts. Second, this result, together with the result that the estimate for θ_3 is -0.0016 and insignificant, show that demand for the treated UPCs during the treatment period in the treatment store is less price elastic relative to the control and the effect is statistically the same for all treated UPCs. This implies that nutritional information signals conveyed through the expert's opinion affect consumer price sensitivity and consumers exposed to such signals are less price sensitive than consumers in the control. Finally, though statistically insignificant, the result that θ_3 is -0.0016 also implies that in the treatment store during the treatment period, UPCs with high NuVal scores and low prices are the favorites of the consumers.

5.2 Robustness Checks

To check the sensitivity of our findings above, we perform several robustness checks. The key identifying assumption in our DID regression analysis above is that trends in the sales of the treated UPCs would be the same in both the treatment store and the control stores during the treatment period in absence of the posting of NuVal scores. This assumption would be violated if, for example, the treatment store repositioned itself through marketing campaigns as the place to shop for a healthier diet and attract more health conscious shoppers around the time when NuVal labels were adopted. This is not implausible as NuVal is a voluntary program. If the trends were not the same for the aforementioned or other reasons, then part of the estimated treatment effect could simply reflect the different sales trends in the treatment and control stores and hence our estimate of the treatment effect would be biased. We perform two robustness checks to examine the validity of this assumption.

In our DID analysis above, we include all grocery stores other than the treatment store in the control group. From Table 2, it is clear that some control stores are more similar to the treatment store than others in terms of the number of treated UPCs offered, the number of units sold per week and the price charged. The advantage of including more control stores is that we will have a larger sample and the parameters of interest will be more precisely estimated. The disadvantage of doing so is that by including those control stores that are not very similar to the treatment store, the key identifying assumption that sales trends are similar in the treatment store and the control stores in absence of treatment may fail. In our first robustness check, we only include one store in the control group. The selected store is control store 2. From Table 2, it is clear that this is the control store that is most similar to the treatment store during the control period in terms of the number of treated UPCs offered, number of units sold per week and price charged. Only specification (2) is estimated and results are reported in Table 7. Compared with the results in Table 6, the only significant change is that the absolute value of the coefficient estimate for the base treatment status variable $D_s * D_t$ is now larger. Also, now it is statistically significant at the 10% level. The estimated treatment effect is,

(6) $-0.1438 + 0.0036 * S_i + 0.2715 * \log p_{ist} - 0.0023 \log p_{ist} * S_i$.

Using the mean value of the logarithm of price for UPCs in the treatment store during the treatment period, the treatment effect becomes a function of the NuVal score only,

(7) $-0.1178 + 0.0034 * S_i$,

which implies that the treatment effect is only positive for those UPCs with a NuVal score that is at least 35. This further strengthens our result that expert opinion matters because not all publicity is good publicity as the treatment effect for UPCs with low NuVal scores is actually negative.

In our second robustness check, we relax the identifying assumption that sales trends in the treatment and control stores would be the same in absence of treatment by including a quadratic store-specific trend in specification (2), that is, terms $\gamma_s t$ and $\gamma_s t^2$ are added to (2). This allows treatment and control stores to follow different trends in a limited but potentially revealing way. Results from this regression are displayed in Table 8. As we can see, again, results are very similar to those reported in Table 6, lending support to the validity of our key identifying assumption.

Our final robustness check is a set of falsification or placebo tests. In each of the placebo test, a single control store is used as a placebo for the treatment store and the remaining four stores are used as the control stores. Both specifications (1) and (2) are estimated for each placebo test and the estimation results are displayed in Table 9. Results show there is no clear and coherent evidence for the false NuVal treatment effect in each of the placebo test, lending support to the supposition that our main results above are not driven by some other unobserved (to the econometrician) advertising or demand shifters. In tests where the first and third control stores are used as the placebo, the estimated treatment effect does not vary with the NuVal score in the sense that the estimated coefficient for the treatment status and NuVal score interaction variable is not statistically significant. In the test where the fifth control store is used as the placebo, the estimated treatment effect decreases with the NuVal score. In the test where the fourth control store is used as the placebo, the estimated overall treatment effect is negative, driven by the large, negative and significant estimate for the coefficient for the treatment status variable. Finally, in the tests where the second and the fifth control stores are used as the placebo, the estimated treatment effect is not robust across different specifications. For both tests, results from (1) indicate there is no statistically significant treatment effect, while results from (2) tell us the opposite.

6 Conclusions

In this paper, we offer a case study on the effect of expert opinion about a credence attribute of the product on consumer demand. As consumers eat many food items every day and a person's health is also influenced by many other factors, it is very difficult for consumers to learn the effect of a particular food item on their health. For the same reason, it is also difficult for the food producers to prove to the consumers the effect of their products on their health. Our results show that in such a scenario, experts are opinion influencers. Consumers purchase more of the products that health and nutrition experts deem healthier.

Our results have important public policy implications. Obesity and nutrition-related chronic diseases impose significant economic and human tolls on society. Medical costs attributable to obesity alone are estimated to be at least $147 billion per year (Finkelstein et al. 2009). Poor diet is a major contributor to obesity and nutrition-related chronic diseases. For example, since the

1960s, medical research has established the effects of excessive intake of saturated fat and cholesterol on heart disease and cancer risks (e.g., Van Horn et al. 2008). There is also convincing evidence that high sodium intake is causally associated with risk for hypertension (IOM 2010). The diet for the majority of the U.S. population does not meet the joint USDA and DHHS Dietary Guidelines for Americans (DGA) (USDA & DHHS 2010). Per capita caloric intake from solid fats and added sugars exceeds the recommended limit by 180%, the highest percentage of foods and food components consumed excessively by Americans, followed by refined grains (100%) and sodium (49%) (USDA & DHHS 2010). Existing nutrition policies have had limited success in improving diet quality at the population level. Although intake of saturated fats has declined over time across demographic subgroups, prevalence of hypertension in the U.S. population stayed at 28 to 30% between 1999 and 2006 (Ostchega et al. 2008) and is estimated to be responsible for 395,000 deaths in the United States annually (Danaei et al. 2009). Our results show that labels like NuVal, which conveys expert opinion, is effective at providing nutrition cues to shoppers and may be effective in promoting healthy food choices at the point of purchase.

Moving forward, there are several questions remain to be answered. First, as we only focus on the 4-month window around the adoption of the NuVal label, our analysis is short-run in nature. In the long run, consumers who shop at the treatment store may learn about the healthfulness of the products through NuVal and pass the information to shoppers at other stores. In this case, even if other stores do not adopt the NuVal label, there will be a spillover effect. Empirically quantifying this effect is challenging because in the long-run, firms introduce and terminate product offerings, which can contaminate the effect of our interest. Second, what is the value of labels with health expert opinion? Knowing how much consumers value such labels can help policy makers decide whether to spend resources to design and promote such labels. Answering this question would require the estimation of a structural model to recover consumer preferences. These are left for future research.

References

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

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

Cabral, L. and A. Hortacsu (2010): "The Dynamics of Seller Reputation: Evidence From eBay," *Journal of Industrial Economics*, 58, 1, 54-78.

Chiuve, S. E., L. Sampson, and W. C. Willett (2011): "The Association between a Nutritional Quality Index and Risk of Chronic Disease." *American Journal of Preventive Medicine, 40*, 505–513.

Danaei, G., E. L. Ding, D. Mozaffarian, B. Taylor, J. Rehm, C. J. L. Murray and M. Ezzati (2009): "The Preventable Causes of Death in the United States: Comparative Risk Assessment of Dietary, Lifestyle, and Metabolic Risk Factors," *PLoS Medicine*, 6, 4.

Darby, M. R. and E. Karni (1973): "Free Competition and the Optimal Amount of Fraud," *Journal of Law and Economics*, 16, 1, 67-88.

Dranove, D. and G. Z. Jin (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48, 4, 935-963.

Drichoutis, A., P. Lazaridis and R. M. Nayga (2006): "Consumers' Use of Nutritional Labels: A Review of Research Studies and Issues," *Academy of Marketing Science Review*, 10, No. 9.

Dube, J.-P., G. J. Hitsch and P. E. Rossi (2010): "State Dependence and Alternative Explanations for Consumer Inertia," *Rand Journal of Economics*, 41, 3, 417-445.

Elfenbein, D., R. Fisman and B. McManus (2012): "Charity as a Substitute for Reputation: Evidence from an Online Marketplace," *Review of Economic Studies*, 79, 1441-1468.

Eliashberg, J. and S. M. Shugan (1997): "Film Critics: Influencers or Predictors?" *Journal of Marketing*, 61, 2, 68-78.

Finkelstein, E. A., J. G. Trogdon, J. W. Cohen and W. Dietz (2009): "Annual Medical Spending Attributable to Obesity: Payer- and Service-specific Estimates." *Health Affairs*, 28, 5, w822–w831.

Fisher, L. M., L. A. Sutherland, L. A. Kaley, T. A. Fox, C. M. Hasler, J. Nobel, M.A. Kantor and J. Blumberg (2011): "Development and Implementation of the Guiding Stars Nutrition Guidance Program," *American Journal of Health Promotion*, 26, 2, 55-63.

Hendel, I. and A. Nevo (2013): "Intertemporal Price Discrimination in Storable Goods Markets," *American Economic Review*, 103, 7, 2722-2751.

Hersey, J. C., K. C. Wohlgenant, J. E. Arsenault, K. M. Kosa and M. K. Muth (2013): “Effects of Front-of-Package and Shelf Nutrition Labeling Systems on Consumers,” *Nutrition Reviews*, 71, 1, 1-14.

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

Imbens, G. W. (2004): “Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review,” *Review of Economics and Statistics*, 86, 4-29.

Institute of Medicine (2010): *A Population-based Policy and Systems Change Approach to Prevent and Control Hypertension*. Washington, DC: The National Academies Press.
Institute of Medicine (2012): *Front-of-package Nutrition Rating Systems and Symbols: Promoting Healthier Choices*. Washington, DC: The National Academies Press.

Ippolito, P. M. and A. D. Mathios (1990): “Information, Advertising and Health Choices: A Study of the Cereal Market,” *Rand Journal of Economics*, 21, 3, 459-480.

Katz, D. L., V. Y. Njike, D. Kennedy, and J. Treu (2007): Overall Nutritional Quality Index, version 1: reference manual. Derby CT: Yale University School of Medicine. http://www.nuval.com/images/upload/file/ONQI%20Manual%205_5_09.pdf

Katz, D. L., V. Y. Njike, L. Q. Rhee, A. Reingold and K. T. Ayoob (2010): “Performance Characteristics of NuVal and the Overall Nutritional Quality Index (ONQI),” *American Journal of Clinical Nutrition*, 91, 1102S–1108S.

Kiesel, K. and S. B. Villas-Boas (2013): “Can Information Costs Affect Consumer Choice? Nutritional Labels in a Supermarket Experiment,” *International Journal of Industrial Organization*, 31, 153-163.

Ollberding, N. J., R. L. Wolf and I. Contento (2010): “Food Label Use and Its Relation to Dietary Intake among US Adults.” *Journal of American Dietetic Association, 110*, 1233–1237.

Ostchega, Y., S. S. Yoon, J. Hughes and T. Louis (2008): *Hypertension Awareness, Treatment, and Control-continued Disparities in Adults: United States, 2005-2006*. NCHS Data Brief No. 3. Hyattsville, MD: National Center for Health Statistics.

Todd, J. E. and J. N. Variyam (2008): “The Decline in Consumer Use of Food Nutrition Labels, 1995-2006.” Economic Research Report Number 63. Washington, DC: Economic Research Service, U.S. Department of Agriculture.

Mathios, A. D. (2000): “The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market,” *Journal of Law and Economics*, 43, 2, 651-677.

Nelson, P. (1970): “Information and Consumer Behavior,” *Journal of Political Economy*, 78, 2, 311-329.

Rahkovsky, I., B.-H. Lin, C.-T. J. Lin and J.-Y. Lee (2013): “Effects of the Guiding Stars Program on Purchases of Ready-to-eat Cereals with Different Nutritional Attributes,” *Food Policy*, 43, 100-107.

Reinstein, D. A. and C. 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, 1, 27-51.

U.S. Department of Agriculture and U.S. Department of Health and Human Services (2010): “Dietary Guidelines for Americans 2010,” Retrieved from www.dietaryguidelines.gov.

Van Horn, L., M. McCoin, P. M. Kris-Etherton, F. Burke, J .A. S. Carson, C. M. Champagne, W. Karmally and G. Sikand (2008): “The Evidence for Dietary Prevention and Treatment of Cardiovascular Disease,” *Journal of the American Dietetic Association*, 108, 287–331.

Wansink, B. and J. Sobal (2007): “Mindless Eating: The 200 Daily Food Decisions We Overlook,” *Environment and Behavior*, 39, 106–123.

Zhu, C. and R. Huang (2014): “Heterogeneity in Consumer Responses to Front-of-Package Nutrition Labels: Evidence from a Natural Experiment,” China Agricultural University Working Paper.

Table 1 Descriptive Statistics for Treated and Other UPCs in the Treatment Store during the Control Period

Variable	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max	Obs
	Treated UPCs					Untreated UPCs				
NuVal Score	51.80	25.58	23	100	130					
Units sold	36.18	49.14	1	679	2290	18.50	38.62	1	711	4076
Unit price	1.43	1.12	0.35	4.79	2290	2.12	1.44	0.38	7.39	4076
Medium size ad	0.01	0.09	0	1	2290	0.01	0.10	0	1	4076
Minor display	0.01	0.11	0	0	2290	0.01	0.09	0	1	4076
General Mills	0.52	0.50	0	1	130	0.20	0.40	0	1	244
Dannon	0.33	0.47	0	1	130	0.32	0.47	0	1	244
Soy	0.03	0.17	0	1	130	0.03	0.17	0	1	244
Greek	0.05	0.23	0	1	130	0.10	0.30	0	1	244
Yogurt Drink	0.05	0.23	0	1	130	0.06	0.23	0	1	244
Organic	0.05	0.21	0	1	130	0.20	0.40	0	1	244

Table 2 Descriptive Statistics for UPCs in Different Stores during the Control Period

Variable	#	NuVal Score	Units Sold	Unit Price	#	Units Sold	Unit Price
	Treated UPCs				Other UPCs		
Treatment Store	130	51.80 (25.58)	36.18 (49.14)	1.43 (1.12)	244	18.50 (38.62)	2.12 (1.44)
Control Store 1	124	52.00 (25.51)	21.82 (33.16)	1.42 (1.06)	193	13.41 (36.38)	1.91 (1.57)
Control Store 2	123	52.61 (25.93)	34.48 (62.53)	1.43 (1.09)	204	18.52 (53.70)	1.99 (1.58)
Control Store 3	118	53.42 (26.14)	26.40 (65.46)	1.46 (1.09)	204	13.15 (29.33)	1.95 (1.57)
Control Store 4	83	48.94 (22.73)	9.00 (8.11)	1.18 (1.00)	92	6.92 (6.94)	1.45 (1.09)
Control Store 5	119	52.94 (25.60)	25.96 (49.62)	1.42 (1.11)	200	10.48 (19.80)	1.89 (1.53)

Notes: Standard deviations are in parentheses.

Table 3 Differences in Means

	Treatment Store	Control Stores	Treatment Store & Score >=50	Control Stores & Score>=50	Treatment Store & Score<50	Control Stores & Score<50
Units (Control Period)	36.18 (49.14)	24.65 (50.91)	51.17 (65.13)	31.20 (61.61)	21.91 (16.06)	17.56 (34.52)
Units (Treatment Period)	45.00 (105.03)	19.34 (37.89)	70.40 (145.26)	24.54 (46.12)	20.83 (16.22)	13.59 (24.71)
Price (Control Period)	1.43 (1.12)	1.40 (1.08)	1.07 (0.86)	1.09 (0.81)	1.77 (1.22)	1.73 (1.23)
Price (Treatment Period)	1.43 (1.11)	1.37 (1.05)	1.06 (0.85)	1.07 (0.78)	1.79 (1.21)	1.70 (1.19)
Number of UPCs	130	130	63	63	67	67

Table 4 Summary Statistics for the Regression Sample

	Mean	Std. Dev.	Min	Max
Log of Units Sold	2.44	1.22	0	7.19
Log of Unit Price	0.08	0.68	-1.27	1.61
Large-size Ad	0.02	0.15	0	1
Medium-size Ad	0.04	0.20	0	1
Minor Display	0.03	0.17	0	1
Major Display	0.01	0.07	0	1
Price Reduction>5%	0.22	0.41	0	1
D_s	0.20	0.40	0	1
D_t	0.50	0.50	0	1
$D_s * D_t$	0.10	0.30	0	1

Note: total number of observations: 23,578.

Table 5 Estimation Results for Specification (1)

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	0.1367***	0.04	0.1371***	0.04
Log of Unit Price	-2.4000***	0.19	-2.1846	0.23
Large-size Ad			-0.1106	0.22
Medium-size Ad			-0.2178**	0.11
Minor Display			0.5694***	0.06
Major Display			0.2964***	0.11
Price Reduction>5%			0.1369***	0.05
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.75		0.76	

Note: ** and *** denote statistical significance at 5% and 1%, respectively.

Table 6: Estimation Results for Specification (2)

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	-0.0445	0.07	-0.0320	0.07
Log of Unit Price	-2.4129***	0.20	-2.2412***	0.23
$D_s * D_t$* Log of Unit Price	0.2466***	0.08	0.2200***	0.08
$D_s * D_t$* Score	0.0031***	0.001	0.0029***	0.001
$D_s * D_t$* Score*Log of Unit Price	-0.0016	0.001	-0.0010	0.001
Large-size Ad			-0.1448	0.21
Medium-size Ad			-0.2455**	0.11
Minor Display			0.5678***	0.06
Major Display			0.3030***	0.11
Price Reduction>5%			0.1236**	0.05
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.75		0.76	

Note: ** and *** denote statistical significance at 5% and 1%, respectively.

Table 7: Robustness Check: Using Only One Control Store

	Estimate	Std. Err.
$D_s * D_t$	-0.1438*	0.07
Log of Unit Price	-2.3743***	0.20
$D_s * D_t$* Log of Unit Price	0.2715***	0.09
$D_s * D_t$* Score	0.0036***	0.001
$D_s * D_t$* Score*Log of Unit Price	-0.0023	0.002
Large-size Ad		
Medium-size Ad		
Minor Display		
Major Display		
Price Reduction>5%		
UPC Fixed Effects	included	
Store Fixed Effects	included	
Week Fixed Effects	included	
Adjusted R-squared	0.81	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Table 8: Robustness Check: Allowing Store-specific Trends

	Estimate	Std. Err.
$D_s * D_t$	-0.0138	0.10
Log of Unit Price	-2.4076***	0.20
$D_s * D_t$* Log of Unit Price	0.2485***	0.08
$D_s * D_t$* Score	0.0031***	0.001
$D_s * D_t$* Score*Log of Unit Price	-0.0016	0.001
Large-size Ad		
Medium-size Ad		
Minor Display		
Major Display		
Price Reduction>5%		
UPC Fixed Effects	included	
Store Fixed Effects	included	
Week Fixed Effects	included	
Adjusted R-squared	0.75	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Table 9a: Results from the Placebo Store 1

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	0.1382***	0.05	0.0009	0.14
Log of Unit Price	-2.4277***	0.21	-2.4473***	0.22
$D_s * D_t$* Log of Unit Price			0.3718**	0.15
$D_s * D_t$* Score			0.0031	0.002
$D_s * D_t$* Score*Log of Unit Price			-0.0012	0.003
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.72		0.73	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Table 9b: Results from the Placebo Store 2

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	0.0458	0.06	-0.0944	0.07
Log of Unit Price	-2.4229***	0.21	-2.4078***	0.21
$D_s * D_t$* Log of Unit Price			0.1358	0.09
$D_s * D_t$* Score			0.0026**	0.001
$D_s * D_t$* Score*Log of Unit Price			-0.0053***	0.002
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.72		0.72	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Table 9c: Results from the Placebo Store 3

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	-0.1138*	0.06	0.0051	0.12
Log of Unit Price	-2.4265***	0.21	-2.4340***	0.21
$D_s * D_t$* Log of Unit Price			0.1235	0.16
$D_s * D_t$* Score			-0.0024	0.002
$D_s * D_t$* Score*Log of Unit Price			-0.0010	0.002
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.72		0.72	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Table 9d: Results from the Placebo Store 4

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	-0.1203*	0.07	-0.3324	0.10
Log of Unit Price	-2.4212***	0.21	-2.4202***	0.21
$D_s * D_t$* Log of Unit Price			0.0810	0.10
$D_s * D_t$* Score			0.0039***	0.001
$D_s * D_t$* Score*Log of Unit Price			-0.0021	0.002
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.72		0.72	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Table 9e: Results from the Placebo Store 5

	Estimate	Std. Err.	Estimate	Std. Err.
$D_s * D_t$	0.0866	0.06	0.2104***	0.08
Log of Unit Price	-2.4223***	0.21	-2.4166***	0.21
$D_s * D_t$* Log of Unit Price			-0.1950***	0.07
$D_s * D_t$* Score			-0.0020*	0.001
$D_s * D_t$* Score*Log of Unit Price			0.0015	0.002
UPC Fixed Effects	included		included	
Store Fixed Effects	included		included	
Week Fixed Effects	included		included	
Adjusted R-squared	0.72		0.72	

Note: *, ** and *** denote statistical significance at 10%, 5% and 1%, respectively.

Figure 1: Examples of Price Tags with NuVal Scores

Note: rating from 1 (the least healthy) to 100 (the healthiest).