The Impact of NuVal Shelf Nutrition Labels on Consumption: Evidence from Cold Cereal Purchases

Grace Melo[a]; Chen Zhen[b]

[a] Ph.D. candidate, Department of Agricultural & Applied Economics, University of Georgia, Athens, GA (gmelo@uga.edu)

[b] Associate Professor, Department of Agricultural & Applied Economics, University of Georgia, Athens, GA (czhen@uga.edu)

Preliminary – Do not cite without authors' permission

Corresponding Author Contact Information:
Grace Melo
Ph.D. Candidate
Department of Agricultural and Applied Economics
The University of Georgia
207 Conner Hall
Athens, GA 30602
Phone: (706) 201-5198
Email: gmelo@uga.edu

December 2016

Abstract

Research examining the effect of summary shelf nutrition labels on consumers' behavior in real market settings is scarce. Using a supermarket's voluntary adoption of NuVal—a 1 to 100 numeric summary shelf label system—as a natural experiment, we estimate a Two-Part Model (TPM) to identify the effect of the NuVal label on consumer purchasing decisions for cold cereal. Our results show that posting the NuVal score not only increases the purchase volume of healthier cold cereal products but also increases households' likelihood to purchase cold cereal products with higher nutrition scores. Tests for heterogeneous treatment effects reveal that lower-income households experience a large improvement in their food choices when the NuVal scores are posted.

Acknowledgment: This research was supported by a grant from Healthy Eating Research, a national program of the Robert Wood Johnson Foundation

Key Words: Shelf Nutrition Labels, NuVal, Two-Part Model

JEL: Q18, D12

Introduction

Because of the proliferation of information about linkages between diet and health, the demand for higher quality foods has been rising. The food and retailing industries have responded to this increase by 1) adapting the composition of foods, 2) offering higher quality options, and 3) engaging in a variety of marketing strategies that signal specific product attributes such as nutritional quality. The Nutritional Facts Panel (NFP)—a government initiative to improve the diets of US consumers—has signaled the nutritional quality of foods since 1994. Nevertheless, more than half of US consumers do not read the NFP (Blitstein and Evans, 2006), and one of the main reasons is that consumers have a poor understanding of nutritional information (Howlett et al., 2008).

Evidence has shown that consumers respond to nutritional information delivered in a simple and concise manner (Van Kleef et al., 2008). This is reflected in the popularity among consumers of simplified nutritional label systems developed by the food industry such as Front-of-Package (FOP) (Derby and Levy, 2001). The proliferation of these labels, however, has made it difficult for consumers to distinguish the nutritional quality of processed foods. As a result, the FDA is considering the development of a standardized nutrition label that provides clear and concise nutritional information on processed foods.

Summary shelf nutritional labels summarize the NFP information in a simple manner; therefore, compared to current labeling systems that are nutrient-specific (e.g., FOP), these labels seem a better alternative to inform consumers of the nutritional value of foods. Evidence regarding whether NuVal, a summary shelf label system, is effective in influencing consumers' behavior at the store is scarce. More importantly, it is not known whether shoppers that have a limited understanding of nutritional information (e.g., low-income households) improve their diets because of the labels. This study investigates the impact of NuVal shelf nutritional labels on consumers' purchase decisions and attempts to determine whether or not the labels are effective in targeting different demographic groups, including low-income shoppers, who are at high risk of obesity.

The most recent study by Zhen and Zheng (2015) reported that NuVal labels increase sales of healthier yogurt products. Nevertheless, because their study used store-label data, they were unable to identify whether the NuVal labels contribute to significant health improvements among shoppers that lack understanding of nutritional information. Nikolova and Inman (2015) evaluated the impact of NuVal on food choices using household-level data and found that the labels improve the nutritional quality of shoppers' food purchases. Yet, they do not assess whether the labels have higher nutrition impacts among different household groups (e.g., low-income vs. high-income households). In addition, their analysis fails to capture the overall impact of the NuVal labels on consumer behavior because their data is contingent on households making purchases. NuVal labels might make households more likely to participate in a purchase of healthier products. Therefore, in this article, we present unique evidence assessing the impact of NuVal shelf nutritional labels on consumer behavior across different household groups. The estimation of heterogeneous treatment effects will allow us to determine whether or not the NuVal labels are targeting the segment of households that can benefit from nutritional advice from health experts when making food choices at the store (e.g., low-income households with poor educational attainment). In lieu of using conditional analysis, we employed a Two-Part Model (TPM) analysis that allows us to explore the effect of NuVal labels on the consumers' likelihood of making a purchase (participation decision) and the quantity purchased (quantity decision). To our knowledge, this study presents the first evidence of the heterogeneous impact of NuVal labels on shoppers' food choices.

In the remainder of this article, we first highlight critical literature focusing on the impact of shelf nutritional labeling and present an overview of different estimation methods to analyze scanner data. Then, we provide a theoretical framework of the demand for health that served as the basis to explain the test for heterogeneous effects in our empirical model. Next, we summarize the data characteristics and describe the regression analysis to assess the impact of the NuVal labels. Finally, we end with conclusions and policy implications.

Shelf Nutrition Labeling and Consumer Choice

Research based on experiments and observational data shows that consumers are interested in learning the nutritional value of foods. This is reflected in consumers' use of simplified nutritional information systems such as FOP labels (Hersey et al., 2013, Kim et al., 2012) and in the non-trivial value shoppers place on these labels (Gracia et al., 2009). However, FOP labels on healthier products may lead shoppers to consume more calories, a phenomenon known as the health halo effect (Wansink and Chandon, 2006), or to overestimate the nutritional value of less healthy products (Kim et al., 2012). This unintended effect might occur because FOP labels only display information that is based on a single or a few nutrients (Berning et al., 2008) rather than information that is based on the overall nutritional profile of the product (Hersey et al., 2013). In addition, consumers may associate descriptors such as *low-fat* with a poor taste (Teisl et al., 2001) and therefore, ignore nutrition information on food items that satisfy hedonistic needs (Balasubramanian and Cole, 2002).

Similar to nutrient-specific FOP labels, shelf nutrition labeling systems display nutritional information in a simple manner, but unlike FOP labels, they summarize nutrition by an interpretive score that is based on a nutrition scoring algorithm (Berryman, 2014, Hersey et al., 2013). Because nutritional scoring systems via shelf labels are relatively new compared with other nutritional labeling systems such as nutrient-specific FOP labels, research examining their effectiveness in promoting healthier choices using household-level data is scarce. In addition, most studies focus on the Guiding Stars label, the second most used shelf nutrition labeling system in U.S. stores after NuVal[i] (Anand, 2016). Guiding Stars uses a four-point summary label to indicate the healthfulness of the product.

Studies based on store-level data found that Guiding Stars increased the sales share of healthy products relative to less-healthy ready-to-eat cereal (RTEC) products after its first implementation in 2006 (Rahkovsky et al., 2013, Sutherland et al., 2010). However, this increase of the sales share of healthy products was only attributed to the decline of purchases of less-healthy products (Cawley et al., 2015). This effect may occur because Guiding Stars uses a 4-point scale, which makes it difficult for the consumer to distinguish nutrition quality between products that earn the same number of stars. In contrast, NuVal system scores foods on a scale from 1 to 100 based on the Overall Nutritional Quality Index (ONQI®) algorithm that profiles the content of more than 30 nutrients and the quality of four nutrition factors (Katz et al., 2010, NuVal, 2012). Similar to Guiding Stars' algorithm, ONQI is based on the Dietary Guidelines for Americans and therefore penalizes nutrients like saturated fat, trans fat, cholesterol, sodium, and sugar, while rewards nutrients like fiber, vitamins, and minerals (NuVal 2012). Figure 1 provides an example of a price tag with a NuVal score.

To date, few studies have examined the ability of NuVal labels to encourage healthier purchases using purchase transaction data. Zhen and Zheng (2015) found that NuVal increases sales of yogurt products that had been assigned NuVal scores, especially among higher-scoring products. However, because their analysis is based on store-level data, it does not allow identification of the heterogeneous impacts of NuVal labels among different demographic groups. The ability to identify heterogeneous treatment effects is important for analysis and design of policies targeting certain population groups that may be at higher risks of obesity.

Second, Nikolova and Inman (2015)[ii] evaluated the impact of NuVal using household scanner data. They found that after NuVal adoption, shoppers switched to higher-scoring products and that shoppers became less sensitive to prices and more to promotions. Nevertheless, their analysis is restricted to households making purchases. In addition, their analysis does not account for unobservable product characteristics that may be correlated with some explanatory variables (e.g., price) in the model.

Store and Household-Level Scanner Data

Store-level scanner data are often used in consumer research partly because 1) these data have been available to researchers earlier and 2) analysis of store-level data is more straightforward than an examination of household-level scanner data. However, these aggregate data fail to capture heterogeneity of consumers' preferences, which can be an important limitation if identifying heterogeneity in purchase behavior is relevant to the study (e.g., a policy that targets specific consumer segments). While household-level data allow one to measure this heterogeneity, unobserved product characteristics can be correlated with retailer marketing decisions including price and advertising that lead to endogeneity issues. To address endogeneity, one can include fixed effects at the UPC level to control for unobserved product attributes that may be correlated with the explanatory variables.

Problems with panelists' participation and compliance are one limitation of household-level scanner data, making households' purchases a poor representation of the choices of all household shopping at the stores captured in the store data. Using scanner panels that require panelists to present a card at the check-out may suffer less from attrition and fatigue-induced underreporting than diary panels in some situations[iii]; however, this does not eliminate the problem (Gupta et al., 1996). Another important feature of household panel data is that zero purchases at UPC level are more frequent in panel data than in-store data. Because information on price and product availability is missing from the household-level data if the household did not purchase the products, conditional demand analysis (i.e., zero purchases are not included) is a common approach. However, choosing data based on the purchase decision can generate sample selection bias. Fortunately, store-level data can be used to fill in this missing product information. Using both household- and store-level data enables us to account for heterogeneity of consumers' preferences. More importantly, it allows us to address price endogeneity and sample selection bias due to exclusion of non-purchases in the estimation analysis.

No-Purchase Option in Scanner Data

According to Briesch et al. (2008), high incidence of zero purchases in panel data can be attributed to non-structural zeros, which can be due to: 1) endogenous factors (e.g., high prices, competitive promotion activities) that cause a product to have zero purchases and 2) a small sample of households that do not purchase all products at all times. Because in the presence of non-structural zeros, there is little information as to whether the household choice is driving zero purchases, it is difficult to determine whether the zeros can be excluded from the estimation (Little and Rubin, 2014). Therefore, to reduce potential bias in parameter estimates, zero purchases should always be included in the regression analysis except when they are caused by product unavailability, also known as structural zeros.

There are a number of approaches to account for non-purchase behavior. According to Strijnev et al. (2004), existing approaches are quite restrictive. First, the logit models such as the multinomial logit and the nested logit impose some restrictions; marketing- and product-related variables have the same relative importance in the household product choice and the no-purchase decision. In addition, it ignores correlations between households' no-purchase and product-volume choice that arise from unobserved characteristics. To account for these correlations between the two household decisions, the translog utility function can be estimated. However, it imposes a restrictive structure on the correlation configuration. In addition, it is a common practice to model the no-purchase outcome as an additional outcome and estimate a multinomial model (Chintagunta, 2002). However, this simplistic approach is more restricted than the nested logit and the translog utility models.

An alternative approach proposed by Strijnev et al. (2004) is more flexible in terms of defining the correlation structure and the influence of product-specific variables on household's no-purchase and product choice decisions. However, their model does not explain the purchase volume decision.

Alternatively, zero purchases and purchases can be modeled using a TPM. The TPM first developed by Cragg (1971) as an extension of the Tobit model became popular when Duan et al. (1984) employed it to model health care expenditures.

We consider a TPM because it can conveniently estimate the extensive purchase decision and the intensive quantity decision using a two-step estimation approach. The TPM can also be jointly estimated, allowing correlation between the two decisions. This estimation approach is also known as the bivariate Sample Selection model (SSM). Yen (2005) proposed a Multivariate-Sample Selection Model (MSSM), which allows correlations between the error terms of multiple product selections and purchase level equations and can be reduced to the Heckman's bivariate SSM (also known as Type 2 Tobit Model) and, with further restrictions, can be simplified to the TPM. Although MSSM performs better based on likelihood ratio tests in Yen's analysis of cigarettes and alcohol demand, it generated the same conclusions (i.e., similar elasticity estimates) as the TPM and SSM models. Therefore, given the scope of our study, we based our analysis on the standard TPM.

Theoretical Framework

Following the demand for health model by Grossman (1972), we define a two-period utility function of a household as follows:

(1) $U = U(\emptyset_0 H_0, \emptyset_1 H_1, Z_0, Z_1)$

where H is the stock of health, $\emptyset$ is the service flow per unit stock, $h = \emptyset H$ is total consumption of health services, and Z is total consumption of non-health related commodities.

The net investment in the health stock can be specified as follows:

(2) $H_1 - H_0 = I_0 - \delta H_0$

where I is gross investment and δ is the rate of depreciation.

Because households produce gross investment in health and the other commodities Z, the production functions can be defined as:

(3) $I = I(D, TH; K)$

(4) $Z = Z(X, T; K)$

where D is the production input diet [iv], X represents the inputs to produce Z. The other inputs TH and T are time inputs, and K is the stock of capital. Because production functions are homogenous of degree 1 in the inputs, the production function of gross investment in health can be defined as:

(5) $I = Dg(t; K)$

where $t = \frac{TH}{M}$ and the marginal products of the inputs are:

(6) $\frac{\partial I}{\partial TH} = \frac{\partial g}{\partial t}$

(7) $\frac{\partial I}{\partial M} = g - \frac{\partial g}{\partial t} t$

The budget and time constraints can be defined as:

(8) $P_0 M_0 + V_0 X_0 + \frac{P_1 M_1 + V_1 X_1}{1+r} = TW + A$

(9) $TW + TL + TH + T = \Omega$

where P and V are the input prices of M and X, respectively; W represents wage, TW is number of working hours, A represents initial assets, and r is the interest rate. In the time constraint, Ω is the total amount of time and TL is the time lost due to illness. We assume that $\frac{\partial TL}{\partial H} < 0$ and $\frac{\partial I}{\partial TH} > 0$.[v]

We can define the full wealth constraint as:

(10) $$W\Omega + A = R$$

Substituting equations (8) and (9) into equation (10), we can define the full wealth constraint as follows:

(11) $$C_0 + CZ_0 + W_0TL_0 + (C_1 + CZ_1 + W_1TL_1)\frac{1}{1+r} = R$$

where the productions costs of producing gross investment and the other commodities are defined as $C = PD + WTH$ and $CZ = VX + WT$

The optimization problem can be solved by finding the equilibrium quantities of gross investment. For this purpose, our objective is to maximize the utility subject to the full wealth constraint:

(12) $$L = U(\emptyset_0H_0, \emptyset_1H_1, Z_0, Z_1) + \lambda(R - (C_0 + CZ_0 + W_0TL_0 + (C_1 + CZ_1 + W_1T_1)\frac{1}{1+r}))$$

The first-order condition (FOC) for gross investment in the initial period is:

(13) $$\frac{\partial U_1}{\partial h_1}\frac{\partial h_1}{\partial H_1}\frac{\partial H_1}{\partial I_0} = \lambda[\frac{dC_0}{dI_0} + W_1(\frac{\frac{\partial TL_1}{\partial H_1}}{\frac{\partial H_1}{\partial I_0}}\frac{1}{1+r})$$

where $\frac{\partial h_1}{\partial H_1} = G_1, \frac{\partial H_1}{\partial I_0} = 1, \frac{dC_0}{dI_0} = \pi_0$, and $\frac{\partial TL_1}{\partial H_1} = -G_1$

Therefore, equation (13) can be written as:

(14) $$\pi_0 = W_1G_1\frac{1}{1+r} + \frac{\partial U_1}{\partial h_1}\left(\frac{1}{\lambda}\right)G_1 = G_1(W_1\frac{1}{1+r} + \frac{\partial U_1}{\partial h_1}\left(\frac{1}{\lambda}\right))$$

where G represents the marginal product of the stock of health when household produces in the healthy days and π_0 marginal cost of gross investment in healthy days in the initial period. Equation (14) indicates that the marginal cost of gross investment equals the present value of marginal benefits.

The optimal gross investment can also be found by minimizing the production cost subject to the production function as follows:

(15) $$L = WTH + PD + \lambda\big(I - Dg(t;K)\big)$$

The FOCs for gross investment are:

(16) $$P\frac{\partial D}{\partial I} = \lambda\left(\frac{\partial D}{\partial I}g + \frac{\partial g}{\partial t}\frac{\partial t}{\partial M}\frac{\partial D}{\partial I}D\right)$$

(17) $$W = \lambda\frac{\partial g}{\partial t}$$

where $\frac{\partial M}{\partial I} = \frac{1}{g - \frac{\partial g}{\partial t}t}$. Substituting for λ from equation (17) into equation (16) we obtain:

(18) $$\frac{P}{g - \frac{\partial g}{\partial t}t} = \frac{W_0}{\frac{\partial g}{\partial t}} = \pi_0$$

Equation (18) indicates that the increase of gross investment from spending an additional dollar on diet equals the increase in gross investment from spending an additional dollar on time.

We can extend the two-period analysis to *n* periods and write equation (14) as:

(19) $$\frac{\pi_{i-1}}{(1+r)^{i-1}} = W_i G_i \frac{1}{(1+r)^i} + \frac{\partial U_i}{\partial h_i}\left(\frac{1}{\lambda}\right)G_i + (1-\delta_i)\pi_i\frac{1}{(1+r)^i}$$

Equation (19) can be arranged as follows:

(20) $$G_i\left(W_i + \frac{\partial U_1}{\partial h_1}\left(\frac{1}{\lambda}\right)(1+r)^i\right) = \pi_{i-1}\left(r - \frac{\pi_i - \pi_{i-1}}{\pi_{i-1}} + \delta_i\frac{\pi_i}{\pi_{i-1}}\right)$$

Equation (20) implies that the value of the marginal product of the stock of health capital must equal the supply price (user cost) of health capital.

To contrast health capital with other human capital forms, we follow Grossman (1972) approach and ignore the consumption of health from now on (i.e., Pure Investment model). Then, equation (20) and the full wealth constraint can be reduced to:

$$(21) \qquad \gamma_i = G_i W_i / \pi_{i-1} = \left(r - \frac{\pi_i - \pi_{i-1}}{\pi_{i-1}} + \delta_i \frac{\pi_i}{\pi_{i-1}} \right)$$

$$(22) \qquad \frac{1}{(1+r)^i} (W_i h_i - \pi_i I_i) + A = R'$$

where $\pi_i I_i = P_i M_i + WTH_i$ holds because of first-degree homogeneity.

Wage Effects

Because $G_i W_i$ is the marginal product of health capital, an increase in the wage rate W_i raises the marginal product value. This implies, the higher the wage rate, the greater the person's value of an increase in healthy days.

Since the wage rate and the demand level of marginal efficiency of (health) capital (MEC) are positively correlated, an increase in the wage rate from W_1 to W_2 shifts to the right the demand curve of MEC. Therefore, if the cost of capital is fixed, the optimal health stock increases from H_1 to H_2 (Figure 2). Although W_i affects demand for health or gross investment of health capital, it does not affect the supply of gross investment. Therefore, an increase in wage will raise the demand for diet.

Education

To determine the effects of education on the demand for health and diet, we calculate the marginal product of human capital K that can be measured by years of formal schooling completed:

(23) $\frac{\partial I}{\partial K} = \frac{\partial I}{\partial TH} * \frac{\partial TH}{\partial K} + \frac{\partial I}{\partial D} * \frac{\partial D}{\partial K}$

(24) $\frac{\partial I}{\partial K} = TH \frac{\partial g\prime}{\partial K} + D \frac{\partial g - tg\prime}{\partial K}$

where g' and $g - tg'$ are the marginal products of diet *D* and time *TH*, respectively. The percentage change in gross investment by every unit change in *K* can be denoted as $r_H = \frac{\partial I}{\partial K} \frac{1}{I}$. Assuming that *K* increases marginal products by the same percentage, we can write:

(25) $r_H = \hat{g} = \hat{g}' = -\hat{\pi}$

where $\hat{\pi}$ is the percentage change in marginal cost and $\hat{g}$ and $\hat{g}'$ are the percentage change in marginal products of the direct inputs (i.e., diet and time, respectively). Because education increases the marginal products of diet and time, it reduces the demand for these inputs to produce a given amount of gross investment. Hence, an increase in education reduces the marginal cost π. Then, with marginal products and wage rate held fixed, an increase in education would raise the marginal efficiency of health capital and shift the MEC to the right (Figure 2). As a result, the demand for health increases from H_1 to H_2. If the price for diet is fixed, the amount of money spent on diet required to produce gross capital investment will increase.

From the calculations above, one would expect that shelf nutritional labeling systems will help high-income and educated shoppers to maximize health by improving their diet. However, because these shoppers probably have a better understanding of the nutritional quality of foods, the impact of shelf nutritional labels might be lower compared with low-income and low-educated individuals. We test for differences in the impact of these labels across demographic groups in the empirical section.

Empirical Framework

Employing data on a grocery retailer's voluntary adoption of NuVal shelf nutrition labels, we test whether posting summary nutrition score on shelf labels improves consumers' food choices.

Conditional One-level Analysis

We estimate a baseline model that is conditional on observing the shopper making non-zero purchases as follows:

(26) $$v_{hitr} = a_i + a_t + a_r + a_h + b_1 P_{itr} + b_2\, Adopt_{itr} + b_3\, Adopt_{itr} * Score_i + M'_{itr}\gamma + \epsilon_{hitr}$$

where v_{hitr} is the purchase volume of UPC i in week *t* by household *h* from retailer *r*. The terms a_i, a_t, a_r, and a_h are product, time, retailer, and household fixed effects, respectively; P_{itr} is the price per unit of volume[vi] of UPC *i* at store *r* in week *t*; $Adopt_{itr}$ is an indicator variable equal to one if retailer *r* had posted NuVal score of UPC *i* in week *t* and zero otherwise; $Score_i$ is the NuVal score of UPC i; and ϵ_{hitr} is the error term.

The term $Adopt_{itr}$ captures the average treatment effect or salience effect of the NuVal label on the treated UPCs at the treatment store. This salience effect only indicates to consumers that the product has received a NuVal score, hence it does not capture the treatment effect of the nutritional information provided by the experts via the NuVal score. The interaction term $Adopt_{itr} * Score_i$ isolates the effect of providing nutritional information from the overall NuVal label effect.

One might be concerned that the control stores and the treatment store could have increased marketing activities during the treatment period[vii]; therefore, we include the vector M_{itr} that controls for advertising and price discounts in store r for UPC i in week t. The time unit is an IRI week that runs from Monday to Sunday.

Unconditional One-level Analysis

One approach to accounting for no-purchase outcomes in a simple manner is to regress the choice variable y_{hitr}, which includes both non-zero purchases and zero purchases, on the set of regressors in equation (26). As discussed in the previous section, this approach imposed several restrictions. Furthermore, inconsistent parameter estimates might result if the non-zero purchases process differ systematically from the no-purchase decisions (Labeaga, 1999).

Store-Level Data Analysis

To test whether analysis of household scanner data generates similar conclusions as in the analysis of store sales data, we estimate an equivalent model of equation (26) using store-level data. This model, which is conditional on observing stores sales, is specified as:

$$V_{itr} = a_i + a_t + a_r + b_1\, P_{itr} + b_2\, Adopt_{itr} + b_3\, Adopt_{itr} * Score_i + M'_{itr}\gamma + \epsilon_{itr} \quad (27)$$

where V_{itr} is sales volume at store *r* in week *t* of UPC *i* and the remaining variables are defined as in equation (26).

Two-Part Model

A suggested approach to accounting for censored purchases (i.e., zero purchases) is estimating a TPM. We estimate the first part of TPM as the participation equation at the UPC-store-household level and the second part as the purchase quantity decision as shown in equation (26). The TPM has been widely used for examining outcomes where there are large proportions of zeros. For example, Duffey et al. (2010) and Haines et al. (1988) used TPMs to estimate censored food demand equations. A TPM[viii] can analyze continuous variables that exhibit a mixed distribution. Specifically, it can model a mixture of a discrete point-mass variable (i.e., all mass at zero) and a continuous random variable (Lachenbruch, 2002).

Because NuVal scores can make households more likely to purchase scored products and especially higher-scoring products, estimating a model that does not account for these choice probabilities (i.e., ignoring the two-step nature of the decision process) may result in biased estimates about the effect of the label on consumers' behavior (Haines et al., 1988). Moreover, our household scanner data consist of a mass of zero purchases in the first part of the distribution (Table 2) followed by right-skewed data (Figure 3). This non-normal distribution of the non-zero purchase data can be accommodated in the conditional part of the TPM.

Following Shonkwiler and Yen (1999) and Tooze et al. (2002), we define the two-part estimation system as follows:

$$y_{it}^* = f(X_{it}, \beta) + u_{2i} + \epsilon_{it} \tag{28}$$

$$d_{it}^* = Z_{it}\alpha + u_{1i} + \epsilon_{it} \qquad \epsilon_{it} \sim N(0, \sigma_e^2 I);$$

$$y_{it} = \begin{cases} q_{it}, & if \;\; d_{it} = 1 \\ 0, & if \;\; d_{it} = 0 \end{cases}$$

where y_{it} and d_{it} are the observed dependent variables, y_{it}^* and d_{it}^* are the corresponding latent variables, q_{it} indicates positive outcomes, X_{it} and Z_{it} are vectors of exogenous covariates, β and α represent the corresponding parameter vectors, and u_{2i} and u_{1i} are random effects. The store and household subscripts (r and h) are suppressed for notational simplicity. The system in equation (28) says that for the first part, a binary dependent variable d_{it} is used to model the probability of observing non-zero purchases ($d_{it} = 1$).

As in Tooze et al. (2002), we estimate the first part of the system with a logit model which is specified as follows:

(29) $$\text{logit}\big(\text{prob}(d_{it} = 1)\big) = logit(p_{it}) = \log\left(\frac{\pi_{it}}{1-\pi_{it}}\right) = Z_{it}\alpha + u_{1i} + \epsilon_{it}$$

The truncated outcome y_{it} represents the volume purchased of UPC *i* at time *t*. Then, conditional on observing purchases $(y_{it} > 0)$, the second part of the system can be represented by a regression model estimated using data on non-zero purchases, as defined in equation (26).

(30) $$\text{E}(Y_{it}|d_{it} = 1\,) = q_{it} = f(X_{it}\beta) + u_{2i} + \epsilon_{it}$$

$$q_{it} \sim N(\mu, \phi\,)$$

where f is a monotone increasing function (e.g., log-normal or log-gamma[ix]) that will make μ approximately Gaussian (i.e., normal) and ϕ is a dispersion parameter.

Heterogeneous Consumers' Responses to NuVal

Health concerns and nutrition knowledge are some of the predictors of label use (Drichoutis et al., 2006, Tooze et al., 2002). Therefore, educated meal planners and those who are more concerned about nutrition are more likely to use nutritional information (Nayga, 1996). For that reason, we expect that improvement of food choices made by households that have a healthy lifestyle (e.g., non-smokers) and households with well-educated members (e.g., with a college degree) will be modest after the adoption of simplified summary nutritional labeling. To test this effect, we allow for heterogeneous responses to the label by including a vector of consumers' characteristics D_h in our TPM equations. For example, heterogeneous effects of the conditional part of the TPM in equation (26) are incorporated as follows:

(31) $v_{hitr} = a_i + a_t + a_r + b_1 P_{itr} + (b_2 + D_h' b_4) Adopt_{itr} + (b_3 + D_h' b_5) Adopt_{itr} * score + M_{itr}' \gamma + \epsilon_{hitr}$

where the parameter vectors b_4 and b_5 indicate whether responses to NuVal vary across demographic groups. The logit model is specified in the same way in equation (31) but with d_{hitr}, instead of v_{hitr}, as the dependent variable.

Data

We use scanner data for cold cereal from the IRI Academic Data Set (Bronnenberg et al., 2008). We focus on the cold cereal category because of its high purchase volume and large variation in NuVal scores (min 10, max 91). IRI scanner data allows us to track household-level purchases and store-level prices and sales in a small Midwestern city before and after the adoption of NuVal labels. In the study town, only one store adopted NuVal (treatment store) and no other stores in the city adopted either Guiding Stars or NuVal labels during our sample period[x].

Household food purchases by panelists at these retailers were automatically captured at the store checkout (i.e., card panelists). This data collection method reduces the incidence of misreported prices and quantities compared with data collected through in-home scanning (e.g., Nielsen Homescan). Another reason we use card panelists is to minimize attrition issues of panel scanner data. Retailer identities and product UPCs for private-label products, withheld from the public-use version of the IRI Academic Data Set, was provided for this research. UPC-level NuVal scores for cold cereal were obtained from NuVal LLC, NuVal's licensing company.

Because our treatment store adopted NuVal in August 2010, we define September 2010 to December 2011 as the adoption period and January 2009 to August 2010 as the control period in our analysis.

Table 2 indicates that our sample consists of 6 grocery stores[xi], 2652 households, and 186 UPCs that were sold during both treatment and control periods[xii]. The NuVal scores for the products in our regression sample range from 10 to 91, with an average of 30.

Differences between Treatment and Control Stores

Before conducting the regression analysis, we first use summary statistics to compare changes in purchases and sales after NuVal adoption without controlling for covariates. Table 1 provides the average weekly quantity purchased at the treatment and control stores. To evaluate whether the household-level data are representative of the store-level data, this table also provides summary statistics of the store data, which comprises products sold in the treatment and control stores during the control and treatment periods. The first two columns of the table report the mean purchases and sales for all treated UPCs, regardless of the NuVal scores for the household- and store-level data. In the remaining columns of Table 1, we investigate whether the treatment effect is different for UPCs with higher and lower NuVal scores.

First, to examine the differences between the control stores and the treatment store, column 1 provides the summary statistics for the cold cereal products purchased and sold in these stores during the control period. The summary statistics for household-level data shows that households purchased more volume of cold cereal per week in the control stores than in the treatment store during the control period (619 vs. 592 grams). However, the summary statics of store-level data indicates that every week the treatment store sold more volume of every UPC than the control stores during the control period (9440 vs. 5768 grams). During the study period, the treatment store carried 96% of treated UPCs of our sample while the control stores carried 99%. Overall, this indicates that although the treatment store did not carry all UPCs in our sample, it was the market leader for cold cereal products in terms of weekly sales in the study town during our sample period. But, the household-level data does not reflect this feature of the data.

The columns of price, advertising, price reduction, and score in Table 1 show the price per gram of product, which is the price paid by the household including retailer discounts at the treatment and control stores during the treatment and control periods, whether the product had a coupon or any other advertising sign (*Ad*), whether the product had a tag indicating a price reduction (*PR*), and the NuVal scores (*Score*) for cold cereal products during the sample period. Although, the treatment and control stores did not carry the same number of UPCs, prices, and NuVal scores were similar during the study period. The prices per 100 grams of product charged by the treatment and the control stores during the control period were \$0.96 and \$0.91, respectively. The mean scores of the UPCs in the treatment store and the control stores were around 29 (i.e., 29.31 and 29.58, respectively). In terms of marketing activities, the control stores made more promotional efforts than the treatment store. The control stores offered more discounts of at least 5% (25% vs. 14%) and advertised more cereal products (11% vs. 2%) than the treatment store during the control period. These differences in means between the control and treatment stores in terms of marketing strategies indicates the need to control for these covariates using a regression analysis.

Differences in Means

Comparing purchases between control and treatment periods at the treatment store for UPCs with higher, lower, and all scores. We observe that there was an increase in purchase volume by 99 grams of UPCs with scores equal or higher than 50[xiii]. As expect, there was a smaller increase in purchase volume of UPCs with scores lower than 50 (i.e., 9 grams). Because 95% of UPCs in the sample were scored less than 50 (See also Figure 4), the overall effect was a slight increase in households' purchases by 12 grams at the treatment store after the NuVal label adoption.

Comparing sales between control and treatment periods at the treatment store, we observe that the treatment store increased sales by 1205 grams of products with scores of at least 50 but it decreased sales by 490 grams of products with scores less than 50. The overall effect was a decrease in sales by 418 grams for all treated UPCs at the treatment store after the label was adopted.

During the same sample period, in the control stores, the average volume sold for each treated UPC with any score value decreased by 676 grams and the average quantity purchased decreased by 40 grams. The estimated effect, measured as the difference between change in means, suggests that posting NuVal labels increased consumer demand for the treated UPCs by 52.38 grams, or about 9% of 592 grams, the average volume purchased in the treatment store before the NuVal labels were adopted.

There was a small reduction in the prices of cold cereal products after the NuVal adoption at the control and treatment stores (i.e., -$0.04 and -$0.07 per 100 grams of product, respectively). Similarly, promotional activities decreased at both the treatment and control stores after the NuVal labels were implemented. Overall, price discounts decreased by 1 percentage point in all stores and advertising decreased by 4 percentage points at the control stores and by 2 percentage points at the treatment store.

A comparison between store- and household-level data shows important differences. First, differences in the means indicate that there was a smaller percentage change in sales volume than in purchase volume after the NuVal labels were adopted at the treatment store compared with the control period (3% vs. 9% of sales and purchase volume, respectively). Second, the store data indicates that the treatment store was the market leader for the cold cereal products in the study city during the 2009-2011 period in terms of sales, while the household-level data indicates that shoppers made more purchases in the control stores than in the treatment store during this period.

While mean comparisons are informative, this approach does not take into account the effects of observed and unobserved factors related to product, store, and household on consumption. We use regression analysis to control for these factors and estimate the NuVal label effect.

Sample Characteristics

Table 2 reports the summary statistics for all variables used in the analysis for cold cereal. The summary statistics indicates that only 0.1% of the shopping trips correspond to purchases of UPC i in week t at store r. On average, households bought about 605 grams and stores sold approximately 6275 grams of each UPC per week at a price of 0.01 dollars per gram. For an average package size of 409 grams, these values of volume correspond to 1.5 and 16 cold cereal units, respectively. On average, 21% of UPCs were advertised and 37% were labeled with a price reduction mark that indicates a price discount of at least 5%.

The summary statistics related to the socio-demographic information of the shoppers in our sample show that 28% are low-income households[xiv], 19% have children, 57% have not attended college, 7% of households have heads with smoking habits, and the average household consists of two members.

Empirical Results

One-level Analysis

Table 3 shows a comparison of the parameter estimates across models. The first column reports the estimation results for the store-level data, columns 2 and 3 report estimates for the one-level purchase analysis, and the last two columns report results for the TPM estimation. Table 3 also reports the estimated treatment effect at the mean NuVal score, the own-price elasticity for cold cereal, and the *Threshold Score*, defined as the cutoff score above which the treatment effect becomes positive. The negative sign of the parameter estimates for *Adopt* in columns 1, 2, and 4 indicate that the estimated treatment effect on sales and purchases is negative for products with scores lower than the *Threshold Score*. As expected, the parameter estimates of the price (*P*) and the marketing variables- price reduction flag (*PR*) and advertising *(Ad)*- are statistically significant and have the expected signs (i.e., negative for the price and positive for the marketing variables) across all models. Overall, the positive signs of the coefficients on *PR* and *Ad*, imply that if a UPC has a price mark-down or an advertising sign (e.g., discount coupons), the amount purchased and sold of this UPC will increase.

The results of the store-level data analysis provide evidence that there were changes in sales after NuVal labels were adopted by the treatment store (i.e., the parameter estimates for *Adopt* and *Adopt*Score* are statistically significant). We find that posting NuVal scores increases sales for cold cereal products with scores higher than 35, the *Threshold Score*. Because the *Threshold Score* is above the mean score, the estimated treatment effect for a UPC at the average NuVal score is negative (-165.37). This indicates that compared with sales in the control stores, sales for a UPC with a score of 30 (i.e., the mean score) decrease by 165 grams in the treatment store with the posting of the NuVal score. Zhen and Zheng (2015) reported an increase in yogurt sales for all NuVal scores in their sample (i.e., scores ranging from 23 to 100) after the NuVal labels were adopted.

Regression results of the one-level purchase analysis show that posting the NuVal scores increases households' purchases of healthier cold cereal products. First, the analysis conditional on households' purchases (columns 3) shows that there is an increase in households' purchases for all NuVal scores. Therefore, the estimated treatment effect at the mean score is positive (44.67). This indicates that compared with purchases in the control stores, purchases of a UPC with a score of 30 increases by 45 grams. The results of the unconditional model (column 2) show that posting the NuVal scores has a positive effect on purchases of UPCs with scores higher than 15, the *Threshold Score*. Because the *Threshold Score* is lower than the mean score, the estimated treatment effect is positive (0.08).

To be able to compare the treatment effects across these three models (store-level data, unconditional, and conditional models), we compare the estimated treatment effects with the effect of a price change. The estimated treatment effect at the mean NuVal score using store-level data (-165.37) is equivalent to a $0.02 price increase for 100 grams of cold cereal. Using household-level data, the estimated treatment effects for the unconditional and conditional analyses (0.08 and 44.67) are equivalent to a price reduction of $0.06 and $0.59 for 100 grams of product, respectively. Overall, the comparison of the treatment effects in dollar terms indicates that the household-level data shows a higher effect of the NuVal labels compared with the store-level data (at least \$0.06 vs. \$0.02). These results support the findings of Table 1, which shows that a larger effect (measured by difference in means) of the NuVal labels is found when using household-level data rather than store-level data.

As discussed before, the unconditional model in our study introduces bias if the decision to purchase and how much to purchase are two different decision-making processes. In addition, the conditional analysis ignores the possibility that NuVal labels affect the likelihood of a shopper making a purchase of a scored product. To address these limitations, we estimate a TPM in the next section.

Two-Part Model

The estimation results of the extensive purchase and the intensive quantity decisions are presented in columns 4 and 5 of Table 3. The scale parameter in the quantity decision equation is 0.14[xv] which is statistically different from 1; therefore, the hypothesis of an exponential distribution for the data is rejected (See Appendix A). The parameter estimates for price (*P*), advertising (*Ad*), and price reduction tag *(PR)* are statistically significant and have the anticipated signs in both purchasing decision models. Similar to the one-level demand models in columns 1 to 3, the equations of the TPM show that the parameter estimates for *Adopt*Score* are positive. For the participation decision (column 4), this indicates that products with higher NuVal scores are more likely to be purchased than UPCs with lower scores. The results of the logit model show that for a one-point increase in the NuVal score of a UPC, the expected change in log odds is 0.01 (or in odds is 1.01). This indicates that for a one-point increase in score, we expect about 1% increase in the odds of making a purchase of a treated UPC.

Because the NuVal labels not only increase the purchase volume of healthier products but also make households more likely to buy products with higher scores, estimating the treatment effect based exclusively on conditional analysis will fail to capture the entire impact of the NuVal labels on consumer decisions. To take into account changes in the purchase probabilities in the estimation of the impact of NuVal scores on purchases, we estimate the unconditional treatment effect. Using the TPM parameter estimates, we estimate the treatment effect (conditional and unconditional) for a UPC that does not have a price mark-down or a coupon (i.e., *PR* and *Ad* take a value of 0) at the average values of score and price (i.e., 30 and \$0.01 per gram of product). Calculations of the conditional and unconditional treatment effects for a log-gamma distributed variable are shown in Appendix B. The estimated conditional and unconditional treatment effects are 0.001 and 0.0001 grams, which are equivalent to a 6% and 8% increase with respect to estimated purchase volume before NuVal adoption. Although the unconditional treatment effect is lower compared with the estimated treatment effects obtained in the unconditional and conditional models (columns 2 and 3), the impact of the NuVal label expressed in percentage changes is not trivial, especially after taking into account changes in the purchasing probabilities due to the labels.

As a whole, the TPM results indicate that NuVal labels not only affect the quantity decision, but they also affect the purchasing intention. Therefore, the approach to explain consumer behavior by estimating a model conditional on purchases will only reflect a partial impact of NuVal labels.

Finally, we compare the own-price elasticity estimate obtained across models in our study with estimates of previous studies. The mean own-price elasticity estimate for the store sales data is smaller than the estimate for market-level sales of instant cereal reported by Jones et al. (1994) (-1.17 vs. -2.4). The unconditional own-price elasticity estimate in the TPM (See Appendix C for calculations) using household-level data is -2.19, which is smaller than the value reported by Nevo (2001). He reported an estimate of -3.4 when estimating a brand-level demand system for RTEC at household level. Differences in the estimates can be attributed to differences in the data and methodology. Overall, our own-price elasticity estimates confirm that cold cereal is a highly price-elastic product and therefore, the adoption of NuVal labels along with price discounts on healthier products may be an important strategy to improve the nutritional quality of consumer choices of cold breakfast cereals.

Heterogeneous Consumers' Responses to NuVal

As the TPM provides a better characterization of consumer behavior, we test heterogeneous effects in the TPM equations. Table 4 shows the estimation results of the TPM for the whole sample (columns 1 and 2) and the subsample of low-income households (columns 3 and 4).

The results in Table 4 for the whole sample indicate that there is heterogeneity in consumer preferences for cold cereal. The results of the first part (participation decision) in column 1 indicate that low-income households, families with heads that do not have any college education, and households with heads who smoke are less likely to purchase cold cereal. The odds of purchasing a UPC of cold cereal every week for these families are about 0.8 times smaller (i.e., 0.78, 0.83, and 0.76, respectively) than the odds of purchasing a UPC for higher-income, college-educated families, and households headed by non-smokers. In contrast, households with children and larger households are more likely to purchase a UPC of cold cereal on a weekly basis. The odds of purchasing a UPC for households with children are about 1.31 times higher compared with households without children.

The results in the quantity decision of the TPM (column 2) related to preferences for cold cereal across all household groups are similar to the results in the participation decision (column 1). The only exception is families with children, who purchase less volume of a UPC of cold cereal than households with children. Overall, the quantity decision model shows that low-income households, families with heads who smoke, and households with heads that have not attended to college buy less cold cereal.

To test for heterogeneous effects, we interact the treatment variables (*Adopt* and *Adopt*Score*) with the demographic variables. The results of the first part of the TPM (column 1) indicates that families with children, households with heads that have smoking habits, and larger families are less likely to purchase cold cereal products with a higher NuVal score compared with families with no children, households who do not have smoking habits, and smaller families. The results for heterogeneous effects in the second part of the TPM show that low-income families, families with children, and smaller households purchase a greater volume of healthier products when NuVal labels are adopted.

The results of the TPM for low-income households are reported in columns 3 and 4 in Table 2. The results for the heterogeneous effects of the participation equation (column 3) for this sub-sample shows that among low-income households, families without any college education and heads who do not smoke are more likely to purchase healthier products than families with at least some college education and households headed by individuals who smoke. The results of the conditional part of the TPM show that low-income families with children and smaller low-income families purchase higher volumes of healthier products compared with low-income families with no children and larger households when NuVal scores are posted.

Conditional Treatment Effects on Purchases across Demographic Groups

Table 5 reports the estimated conditional treatment effects on purchases across household groups. We report the treatment effects only for those household groups with a sample size greater than 1%[xvi]. Column 1 indicates the proportion of each household group in the sample. Columns 2 to 4 indicate the treatment effects (expressed in percentage change in volume with respect to the predicted purchase volume before NuVal adoption) for UPCs with three different values of NuVal scores (min 10, mean 30, max 91).

The conditional analysis indicates that the treatment effects for a UPC with the minimum score value (i.e., 10) are negative for all shoppers. The negative sign of the treatment effects indicates that shoppers decreased the purchase volume of this UPC after NuVal scores were assigned. The largest decrease in purchases was experienced by low-income families with college-educated household heads who have children and have no smoking habits. This household group decreased purchases of this UPC (i.e., a product with a NuVal score of 10) by 36% due to the NuVal labels. They also experience the largest conditional treatment effect when a UPC with the highest score (i.e., 91) is assigned the NuVal labels. They increased purchases of this UPC by about 59% after the NuVal adoption. The second largest improvement of the healthfulness of food choices occurred among low-income shoppers with children who did not attend college and do not smoke. They increased purchases of the healthiest UPC by 49% and decreased purchases of the unhealthiest UPC by 31%.

In general terms, the category of shoppers that experienced at least 49% increase in purchase volume of products with the highest score and more than 30% decrease of products with the lowest score were low-income households with children with no smoking behavior.

Unconditional Treatment Effects on Purchases across Demographic Groups

While the conditional treatment effects are informative, they do not account for changes in shoppers' likelihood to purchase UPCs with higher and lower NuVal scores. The estimated unconditional treatment effects reported in Table 6 indicate to some extent different results compared with the conditional treatment effects in Table 5. Differences in the estimated treatment effects between the conditional and unconditional analysis are caused because the parameter estimates for the heterogeneous treatment effects related to the four main household groups (i.e., *Low Income*, *No College*, *No Children,* and *Smoke*) in the logit model are negative, while in the quantity model the corresponding parameter estimates are positive (See Table 4).

We find that for a UPC with the lowest score, the largest decrease in purchases (39%) was experienced among low-income families with no college-educated household heads who have no children and reported non-smoking behavior. This household group also experienced the second largest increase in purchases (76%) for a UPC with the highest score. The largest increase in purchases (156%) for this UPC (i.e., with the highest score) was experienced among low-income families, with household heads who have a college education, no children, and do not smoke.

While the results of the conditional analysis indicate that the largest increase in healthier products was among low-income families with college-educated household heads who have children and have smoking habits experienced the largest effect, the results of the unconditional analysis point out a different household group (i.e., low-income families with household heads who have a college education, no children, and do not smoke). Yet, both analyses agree that the adoption of the NuVal labels improved food choices of low-income shoppers at the treatment store in this small Midwestern town.

Conclusions and Implications

With the intention of the FDA to develop a standardized, science-based criteria symbol that provides clear and concise nutritional information of processed foods, examining the effectiveness of current summary nutrition labeling systems on improving food choices is critical. Using a supermarket's voluntary adoption of NuVal— a 1 to 100 numeric summary shelf label system—as a natural experiment, we estimate a Two-Part Model (TPM) to identify the effect of the NuVal label on consumer purchasing decisions for cold cereal. In addition, to test whether the households' purchases are representative of sales at the stores in our sample, we estimate a Difference-in-Difference (DID) model using store-level data. Our main findings are as follows:

First, the results of the DID model indicate that the adoption of the NuVal labels increases sales of healthier cold cereal products. However, the estimated treatment effects between store- and household-level data analysis were to some extent different. Differences in the treatment effects can be due to the differences between sales and purchases shown in the summary statistics for our study period. These differences might provide evidence of self-selection bias in household scanner data.

Second, the results of the purchasing and quantity decisions of the TPM show that posting the NuVal labels not only makes a household buy more units of healthier cold cereal products, but it also increases the probability of a household buying healthier products. Therefore, assessing shoppers' choices based on conditional purchases will fail to capture the overall impact of the NuVal labels on food choices.

Finally, tests for heterogeneous preferences in the TPM indicate that lower-income households experience the largest increase in purchases of healthier cold cereal products when a simplified nutrition label format is introduced. This can be due to the fact that their gap of nutritional information *ex-ante* is larger compared with other household groups. As a whole, our findings suggest that providing interpretative summary nutrition information about the overall nutrition value of food products can be an effective way to improve consumer choices.

Tables

Table 1 Difference between Control and Treatment Stores

		Volume			Volume & Score<50 (176 UPCs)			Volume & Score>=50 (10 UPCs)			Price (100 grams)		Ad		PR		Score
*Period		Control	Trt.	Change	Control	Trt.	Change	Control	Trt.	Change	Control	Trt.	Control	Trt.	Control	Trt.	
<u>Household-Level Data</u>																	
Control Stores	Mean	618.67	578.22	-40.45	618.34	578.75	-39.59	635.65	547.13	-88.52	0.96	0.93	0.15	0.11	0.25	0.24	29.58
(185 UPCs)	S.D.	434.97	374.26		434.19	371.84		473.82	496.26		1.02	0.33	0.36	0.31	0.43	0.43	14.36
Treatment Store	Mean	592.31	604.24	11.94	593.18	602.13	8.95	565.51	664.96	99.45	0.91	0.84	0.05	0.02	0.14	0.12	29.31
(178 UPCs)	S.D.	462.48	523.62		467.01	507.41		291.54	869.78		0.99	0.38	0.22	0.15	0.34	0.33	13.85
Effect (grams)				52.38			48.54			187.97							
Effect (%)				8.84%			8.18%			33.23%							
<u>Store-level Data</u>																	
Control Stores	Mean	5768.34	5092.43	-675.91	5908.14	5254.49	-653.65	2445.97	1570.50	-875.47	0.96	0.93	0.15	0.11	0.25	0.24	29.58
(185 UPCs)	S.D.	15945.07	14668.49		16236.09	14974.30		4492.78	2279.28		1.02	0.33	0.36	0.31	0.43	0.43	14.36
Treatment Store	Mean	9439.66	9021.36	-418.30	9590.04	9099.87	-490.17	5825.27	7030.38	1205.11	0.91	0.84	0.05	0.02	0.14	0.12	29.31
(178 UPCs)	S.D.	29103.46	26698.28		29620.91	26779.88		10163.66	24501.47		0.99	0.38	0.22	0.15	0.34	0.33	13.85
Effect (grams)				257.61			163.48			2080.58							
Effect (%)				2.73%			1.70%			35.72%							

Note: S.D. indicates standard deviation, Trt. indicates Treatment, and Diff. denotes Difference. * Because UPCs were assigned scores on different dates, UPCs scored after the NuVal adoption have unique treatment and control periods.

Table 2 Summary Statistics for the Regression Analysis

Variables	Description	Mean	S.D.
<u>Dependent Variables</u>			
V	Sales volume, number of equivalized units sold of UPC *i* in time *t* from retailer *r*	6274.67	18764.12
Y	Y=Purchase volume if D=1, 0 otherwise	0.87	28.18
D	D=1 if UPC *i* was purchased in time *t* from retailer *r* by household *h*, 0 otherwise	0.00	0.00
v	Purchase volume, Number of equivalized units purchased of UPC *i* in time *t* from retailer *r* by household *h*	605.37	435.08
<u>Explanatory Variables</u>			
Treatment Variables			
Adopt	Adopt=1 if UPC *i* had been assigned a score during the treatment period at the treatment store at time t, 0 otherwise	0.08	0.26
Score	NuVal Score of UPC *i*	29.60	14.45
Marketing variables			
P	Price per equivalized unit	0.01	0.01
Ad	Ad =1 if coupon or if any advertising sign, 0 otherwise	0.21	0.40
PR	Price Reduction flag= 1 if Total Price Reduction is 5% or greater, 0 otherwise	0.37	0.48
Household Characteristics (Household-level Data)			
Low Inc	Low Inc=1 if low-income household according to the FPG, 0 otherwise	0.29	0.45
Children	Children=1 if household has children, 0 otherwise	0.19	0.39
No College	No College=1 if household heads have not attended college, 0 otherwise	0.57	0.50
Smoke	Smoke=1 if both household heads smoke, 0 otherwise	0.07	0.25
HHsize	Household Size	2.34	1.24
UPCs		186	
Households		2652	
Stores		6	
Weeks (2009-2011)		155	

Note: S.D. indicates standard deviation

Table 3 Comparison of Estimation Results

	Store-level Data Analysis			One-Level Purchase Analysis						TPM Purchase Analysis					
	1			2			3			4			5		
				Unconditional Analysis			Conditional Analysis			Participation Decision			Quantity Decision		
	Normal Distribution			Normal Distribution			Normal Distribution			(Logit Model)			(Log-gamma Distribution)		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
P	-733703.73	***	31008.74	-135.29	***	2.60	-7623.87	***	816.54	-243.10	***	3.57	-14.14	***	0.84
Adopt	-1085.54	*	583.11	-0.08	*	0.04	12.60		18.42	-0.32	***	0.05	0.06	***	0.02
Adopt*Score	30.67	*	17.92	0.01	***	0.001	1.07	*	0.58	0.01	***	0.00	0.001		0.001
Ad	8207.95	***	188.82	1.57	***	0.02	33.95	***	4.05	0.43	***	0.01	0.04	***	0.004
PR	7869.74	***	156.33	1.20	***	0.01	29.72	***	4.43	0.78	***	0.01	0.03	***	0.005
Scale													0.14		0.001
Treatment Effect (grams)	-165.37			0.08			44.67			-0.000008			0.411		0.0001[a]
Threshold Score	35			15			1			32			1		22.000[a]
Own-Price Elasticity	-1.17			-1.56			-0.13						-0.12		-2.19[a]
Week FE	yes			yes			yes			yes			yes		
UPC FE	yes			yes			yes			yes			yes		
Store FE	yes			yes			yes			yes			yes		
Household FE				yes			yes			yes			yes		
AIC										872523			870116.8		
R2	0.174			0.003			0.349								
N	116536			44967263			64321			44967263			64321		

Note: S.E. denotes standard error. *** p<0.01, ** p<0.05, * p<0.1. T.E. denotes Treatment Effect. [a] Unconditional Treatment Effects

Table 4 TPM Estimation Results with Heterogeneous Effects

	Whole Sample						Low-income Sample					
	Participation Decision Logit Model			Quantity Decision (Log-gamma Distribution)			Participation Decision Logit Model			Quantity Decision (Log-gamma Distribution)		
Variable	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Intercept	-5.692	***	0.373	5.682	***	0.131	-6.047	***	1.510	5.656	***	0.262
P	-243.100	***	3.575	-16.511	***	0.934	-224.800	***	6.885	-19.520	***	1.849
Adopt	-0.763	***	0.122	-0.148	**	0.063	-0.047		0.254	-0.379	***	0.121
Adopt*Score	0.025	***	0.004	0.005	**	0.002	0.011		0.009	0.012	***	0.004
Low Inc	-0.122	***	0.009	-0.009	**	0.004						
Children	0.274	***	0.013	-0.039	***	0.006	0.237	***	0.025	-0.117	***	0.011
No College	-0.181	***	0.009	-0.007	*	0.004	-0.284	***	0.017	0.017	**	0.008
Smoke	-0.270	***	0.018	-0.023	***	0.008	-0.381	***	0.034	-0.101	***	0.015
HHsize	0.154	***	0.004	0.017	***	0.002	0.134	***	0.007	0.043	***	0.003
Adopt*Low Inc	0.095		0.095	-0.229	***	0.046						
Adopt*Children	0.289	**	0.139	-0.187	***	0.072	-0.053		0.329	-0.372	**	0.148
Adopt*No College	-0.140	*	0.084	0.090	**	0.043	-0.485	**	0.190	-0.060		0.096
Adopt*Smoke	0.307		0.367	-0.044		0.263	1.766	**	0.888	-0.300		0.673
Adopt*HHsize	0.156	***	0.046	0.113	***	0.024	0.084		0.085	0.179	***	0.038
Adopt*Score*Low Inc	-0.002		0.003	0.009	***	0.002						
Adopt*Score*Children	-0.015	***	0.005	0.006	**	0.003	-0.017		0.012	0.012	**	0.005
Adopt*Score*No College	-0.002		0.003	-0.002		0.001	0.012	*	0.006	0.005		0.003
Adopt*Score*Smoke	-0.025	*	0.013	0.001		0.010	-0.078	**	0.035	0.014		0.027
Adopt*Score*HHsize	-0.004	**	0.002	-0.003	***	0.001	-0.004		0.003	-0.005	***	0.001
Ad	0.430	***	0.012	0.070	***	0.005	0.478	***	0.023	0.053	***	0.009
PR	0.779	***	0.013	0.082	***	0.005	0.801	***	0.025	0.063	***	0.010
Scale				0.191		0.001				0.189		0.002
Week FE	yes			yes			yes			yes		
UPC FE	yes			yes			yes			yes		
Store FE	yes			yes			yes			yes		
AIC	872411.51			256174.6			230585.78			233311.1		
N	44967263			19626			12574113			16946		

Note: S.E. denotes standard error. *** p<0.01, ** p<0.05, * p<0.1

Table 5 Conditional Treatment Effects on Purchases across Demographic Groups

Household Group	% Sample	Treatment Effects		
		Min Score	Mean Score	Max Score
Low income, children, college, no smoke	3.51%	-36.28%	-20.12%	59.14%
Low income, children, no college, no smoke	2.71%	-31.43%	-16.90%	49.34%
Low income, no children, college, no smoke	5.20%	-27.71%	-19.80%	10.09%
Low income, no children, no college, smoke	1.55%	-25.09%	-18.57%	5.05%
Low income, no children, no college, no smoke	15.84%	-22.21%	-16.56%	3.32%
High income, children, college, no smoke	8.60%	-26.54%	-22.52%	-8.87%
High income, children, no college, no smoke	3.21%	-20.95%	-19.40%	-14.48%
High income, no children, college, no smoke	23.19%	-16.66%	-22.21%	-36.95%
High income, no children, no college, smoke	3.17%	-13.64%	-21.01%	-39.84%
High income, no children, no college, no smoke	30.81%	-10.32%	-19.07%	-40.83%

Note: Treatment Effects are estimated at the average price, Household size=3, Ad=0, PR=0. The % change is with respect to the predicted purchases before NuVal adoption

Table 6 Unconditional Treatment Effects on Purchases across Demographic Groups

Household Group	% Sample	Treatment Effects		
		Min Score	Mean Score	Max Score
Low income, no children, college, no smoke	5.20%	-33.64%	-7.36%	156.12%
Low income, no children, no college, no smoke	15.84%	-39.07%	-20.83%	75.89%
High income, no children, college, no smoke	23.19%	-29.15%	-13.82%	56.59%
Low income, children, college, no smoke	3.51%	-32.46%	-20.28%	32.20%
High income, no children, no college, no smoke	30.81%	-34.95%	-26.35%	7.55%
Low income, children, no college, no smoke	2.71%	-37.99%	-31.87%	-9.24%
High income, children, college, no smoke	8.60%	-27.90%	-25.83%	-19.16%
High income, children, no college, no smoke	3.21%	-33.80%	-36.62%	-44.50%
Low income, no children, no college, smoke	1.55%	-37.64%	-49.79%	-74.07%
High income, no children, no college, smoke	3.17%	-33.43%	-53.29%	-84.14%

Note: Treatment Effects are estimated at the average price, Household size=3, Ad=0, PR=0. The % change is with respect to the predicted purchases before NuVal adoption

Figures

Figure 1 Example of a Price Tag with NuVal Score for Cold Cereal at a NuVal store

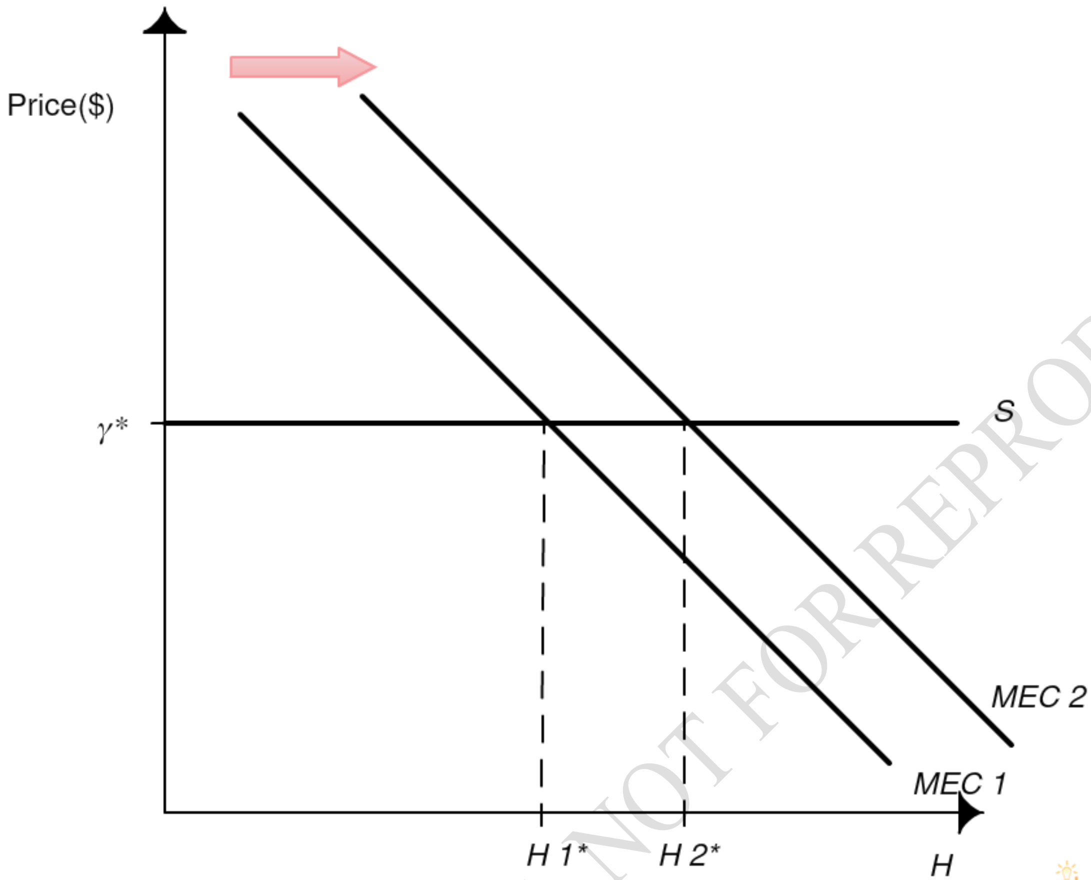

Figure 2 The Marginal Efficiency of Health Capital (MEC) Demand Curve

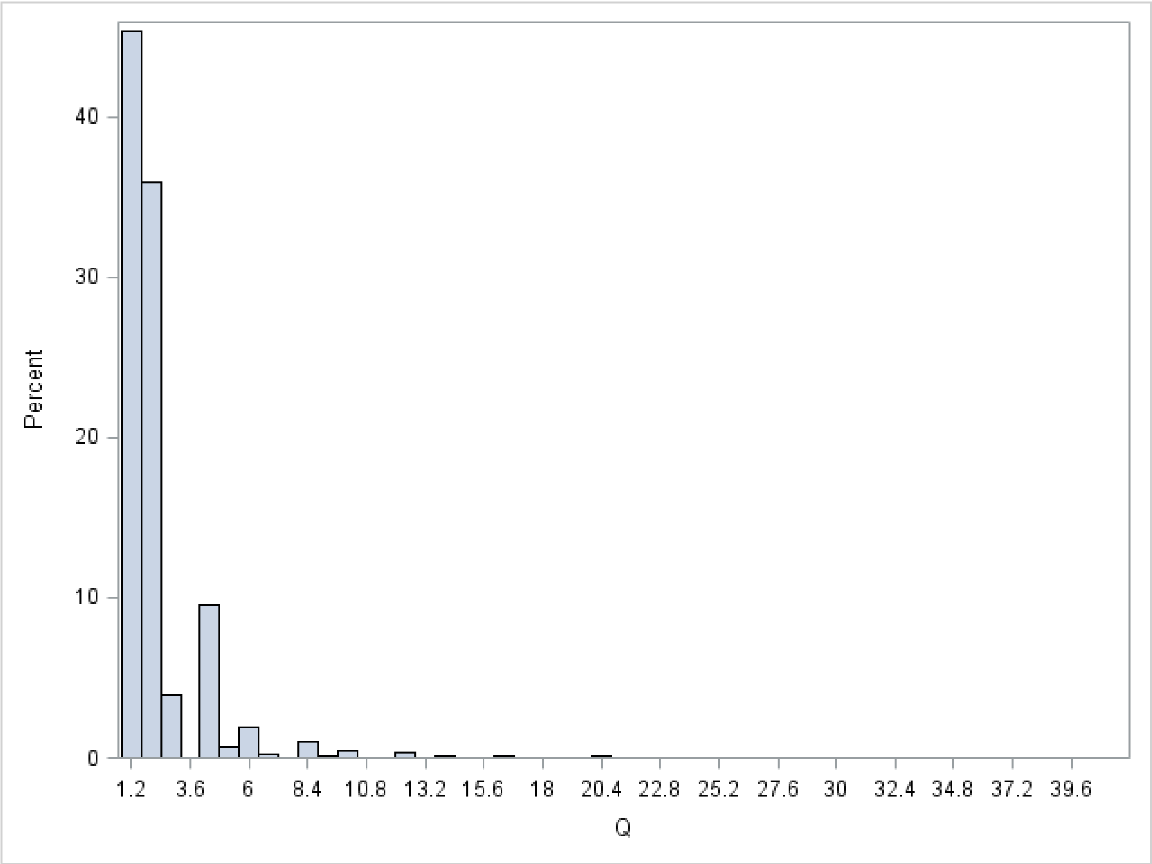

Figure 3 Distribution of Weekly Purchases of Cold Cereal during the Sample Period.

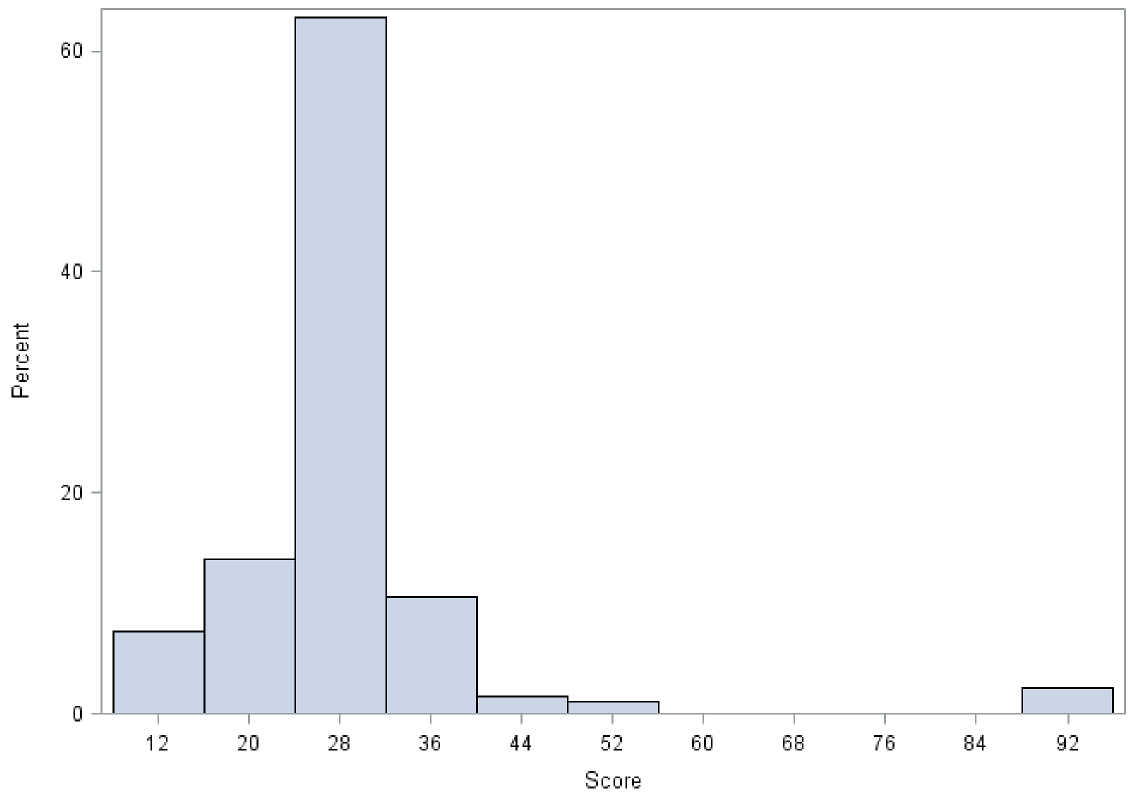

Figure 4 Distribution of NuVal scores of the UPCs in the sample.

Appendix

Appendix A Log-gamma Distribution for the second part of the TPM

As shown in figure 3, the distribution of the purchasing data appears to be skewed to the right. On average, 99.9% of the sample are zero purchases of UPC *i* during a shopping trip to store *r* made by household *h* at week *t* (Table 2). The gamma distribution allows estimating a general form for continuous outcomes that has the form of a peak close to zero or no peak (e.g., the negative exponential) and is decreasing from zero (Simpson et al., 2004).

Let q be a random variable following a generalized log-gamma distribution, which provides tests of the Weibull and log-normal models and includes both the log-normal and the gamma with log link (Manning et al., 2005). Then, according to Manning et al. (2005), the probability density function for the generalized gamma is given by

$$(32) \qquad f(q|\mu,\sigma,k) = \frac{\gamma^{\gamma}}{\sigma y\sqrt{\gamma}\Gamma(\gamma)} exp(z\sqrt{\gamma} - u),\ y \geq 0,$$

Where $\gamma = |k|^{-2}, z = \frac{sign(k)(\ln(y)-\mu)}{\sigma}$, and $u = \gamma\exp(|k|z)$. The parameters $\mu, \sigma,$ and k correspond to position, scale, and shape, respectively. The scale parameter is the inverse of the dispersion parameter ϕ in equation (30). For $k > 0$ the probability density function of q is skewed to the right, while for k <0 q is skewed to the left. A normal distribution of the probability density function is represented by $k = 0$. A value of 1 of the scale parameter corresponds to the exponential distribution.

In the case that $\sigma = k$, the generalized gamma distribution can be reduced to:

$$(33) \qquad f(q|\mu,\sigma) = \frac{\gamma^{\gamma}}{y\Gamma(\gamma)} exp(z\sqrt{\gamma} - \gamma\exp(\sigma z)),\ \sigma > 0$$

Appendix B Unconditional Own-Price Elasticity

According to Yen (2005), the elasticity of the unconditional mean with respect to *a* common element of x and z (say $x_j = z_j$) can be computed by differentiating the unconditional mean $E(y|x) = \Phi(x'\alpha) * \exp(x'\beta + \frac{1}{2}\sigma^2)$.

(1) $$\frac{dE(y|x)}{dx_j} = \exp[x'\beta + \ln(k^2)] * \left(\beta_j * \Phi(x'\alpha) + \alpha_j * \phi(x'\alpha)\right)$$

(2) $$e^u = \frac{dE(y|x)}{dx_j} * \frac{x_j}{E(y|x)} = \left(\exp[x'\beta + \ln(k^2)]\left(\beta_j * \Phi(x'\alpha) + \alpha_j * \phi(x'\alpha)\right)\right) * \frac{x_j}{\Phi(x'\alpha) * \exp[x'\beta + \ln(k^2)]}$$

(3) $$e^u = \left(\beta_j + \alpha_j * \lambda(x'\alpha)\right) * x_j$$

where $\lambda(z'\alpha)$ is the inverse mills ratio. It is clear from the above that the sign of the elasticity of the unconditional mean is negative as long as the own-price elasticity parameter estimates of the purchasing and quantity equations α_j and b_j have negative sign.

Appendix C Treatment Effects

The unconditional predicted purchases can be derived by using the TPM estimates as follows:

(4) $E(y|x) = \Pr(y > 0) * E(y|y > 0, x) = \phi(x'\alpha) * E(y|y > 0, x)$

Where ϕ represents the standard normal cumulative distribution function or logit:

(5) $\Pr(Q > 0) = \frac{\exp(x\prime\alpha)}{\exp(x\prime\alpha)+1}$

The expected value of *y* conditional on *y>0* for a model based on a generalized gamma distribution and a log link relationship (i.e., $\ln(E(y|y > 0, x) = x\beta)$ *is* given by:

(6) $E(y\,|\,y > 0, x) = \exp[x'\beta + \left(\frac{\sigma}{k}\right)\ln(k^2) + \ln\left(\Gamma\left(\left(\frac{1}{k^2}\right) + \left(\frac{\sigma}{k}\right)\right)\right) - \ln\Gamma\left(\left(\frac{1}{k^2}\right)\right)]$

When σ= k (i.e., standard gamma distribution), we can reduce the conditional expectation to:

(7) $E(y\,|\,y > 0, x) = \exp[x'\beta + \ln(k^2)]$

Therefore, the unconditional mean of *y* is defined by:

$$E(y|x) = \frac{\exp(x'\alpha)}{\exp(x'\alpha) + 1} * \exp[x'\beta + \ln(k^2)]$$

And the unconditional treatment effect *TE* can be defined as:

(8) $TE = E(y|\bar{x}_{-k},\ adopt = 1\,) - E(y|, \bar{x}_{-k}, adopt = 0)$

where *adopt* correspond to the treatment variable.

References

Anand, P. (2016) "How to Improve Nutrition Labeling on Food." In *The Wallstreet Journal*.

Balasubramanian, S.K., and C. Cole. 2002. "Consumers' search and use of nutrition information: The challenge and promise of the nutrition labeling and education act." *Journal Of Marketing* 66:112-127.

Berning, J.P., H.H. Chouinard, and J.J. McCluskey. 2008. "Consumer preferences for detailed versus summary formats of nutrition information on grocery store shelf labels." *Journal of Agricultural & Food Industrial Organization* 6.

Berryman, P. 2014. *Advances in Food and Beverage Labelling: Information and Regulations*: Elsevier.

Blitstein, J.L., and W.D. Evans. 2006. "Use of nutrition facts panels among adults who make household food purchasing decisions." *Journal of Nutrition Education and Behavior* 38:360-364.

Briesch, R.A., W.R. Dillon, and R.C. Blattberg. 2008. "Treating zero brand sales observations in choice model estimation: Consequences and potential remedies." *Journal of Marketing Research* 45:618-632.

Bronnenberg, B.J., M.W. Kruger, and C.F. Mela. 2008. "Database paper-The IRI marketing data set." *Marketing science* 27:745-748.

Cawley, J., et al. 2015. "The impact of a supermarket nutrition rating system on purchases of nutritious and less nutritious foods." *Public health nutrition* 18:8-14.

Chintagunta, P.K. 2002. "Investigating category pricing behavior at a retail chain." *Journal of Marketing Research* 39:141-154.

Cragg, J.G. 1971. "Some statistical models for limited dependent variables with application to the demand for durable goods." *Econometrica: Journal of the Econometric Society*:829-844.

Derby, B.M., and A.S. Levy. 2001. "Do Food Labels Work?" *Handbook Of Marketing And Society*:372-383.

Drichoutis, A.C., P. Lazaridis, and R.M. Nayga Jr. 2006. "Consumers' use of nutritional labels: a review of research studies and issues." *Academy Of Marketing Science Review* 2006:1.

Duan, N., et al. 1984. "Choosing between the sample-selection model and the multi-part model." *Journal of Business & Economic Statistics* 2:283-289.

Duffey, K.J., et al. 2010. "Food price and diet and health outcomes: 20 years of the CARDIA Study." *Archives of internal medicine* 170:420-426.

Gracia, A., M.L. Loureiro, and R.M. Nayga. 2009. "Consumers' valuation of nutritional information: a choice experiment study." *Food Quality and Preference* 20:463-471.

Grossman, M. 1972. "On the concept of health capital and the demand for health." *Journal of Political economy* 80:223-255.

Gupta, S., et al. 1996. "Do household scanner data provide representative inferences from brand choices: A comparison with store data." *Journal of Marketing Research*:383-398.

Haines, P.S., B.M. Popkin, and D.K. Guilkey. 1988. "Modeling food consumption decisions as a two-step process." *American Journal of Agricultural Economics* 70:543-552.

Hersey, J.C., et al. 2013. "Effects of front-of-package and shelf nutrition labeling systems on consumers." *Nutrition reviews* 71:1-14.

Howlett, E., S. Burton, and J. Kozup. 2008. "How modification of the nutrition facts panel influences consumers at risk for heart disease: the case of trans fat." *Journal of Public Policy & Marketing* 27:83-97.

Jones, E., W.S. Chern, and B.K. Mustiful. 1994. "Are lower-income shoppers as price sensitive as higher income ones? A look at breakfast cereals." *Journal of Food Distribution Research* 25:82-92.

Katz, D.L., et al. 2010. "Performance characteristics of NuVal and the overall nutritional quality index (ONQI)." *The American journal of clinical nutrition* 91:1102S-1108S.

Kim, H., et al. 2012. "Front-of-package nutritional labels and consumer beverage perceptions." *Applied Economic Perspectives and Policy*:pps037.

Labeaga, J.M. 1999. "A double-hurdle rational addiction model with heterogeneity: estimating the demand for tobacco." *Journal Of Econometrics* 93:49-72.

Lachenbruch, P.A. 2002. "Analysis of data with excess zeros." *Statistical Methods In Medical Research* 11:297-302.

Little, R.J., and D.B. Rubin. 2014. *Statistical analysis with missing data*: John Wiley & Sons.

Manning, W.G., A. Basu, and J. Mullahy. 2005. "Generalized modeling approaches to risk adjustment of skewed outcomes data." *Journal of health economics* 24:465-488.

Nayga, R.M. 1996. "Determinants of consumers' use of nutritional information on food packages." *Journal of Agricultural and Applied Economics* 28:303-312.

Nevo, A. 2001. "Measuring market power in the ready-to-eat cereal industry." *Econometrica* 69:307-342.

Nikolova, H.D., and J.J. Inman. 2015. "Healthy Choice: The Effect of Simplified Point-of-Sale Nutritional Information on Consumer Food Choice Behavior." *Journal of Marketing Research* 52:817-835.

NuVal, L. (2012) "ONQI: The Science Behind The Scores." In.

Rahkovsky, I., et al. 2013. "Effects of the Guiding Stars Program on purchases of ready-to-eat cereals with different nutritional attributes." *Food Policy* 43:100-107.

Shonkwiler, J.S., and S.T. Yen. 1999. "Two-step estimation of a censored system of equations." *American Journal of Agricultural Economics* 81:972-982.

Simpson, P., et al. 2004. "Assessing model fit and finding a fit model." *Sas Users Group International* 29:214-229.

Strijnev, A., P. Seetharaman, and S. Chib. 2004. "Model of Brand Choice with a No-Purchase Option Calibrated to Scanner-Panel Data." *Journal of Marketing Research*:184-196.

Sutherland, L.A., L.A. Kaley, and L. Fischer. 2010. "Guiding stars: the effect of a nutrition navigation program on consumer purchases at the supermarket." *The American journal of clinical nutrition* 91:1090S-1094S.

Teisl, M.F., N.E. Bockstael, and A. Levy. 2001. "Measuring the welfare effects of nutrition information." *American Journal of Agricultural Economics* 83:133-149.
Tooze, J.A., G.K. Grunwald, and R.H. Jones. 2002. "Analysis of repeated measures data with clumping at zero." *Statistical Methods In Medical Research* 11:341-355.
Van Kleef, E., et al. 2008. "Consumer preferences for front-of-pack calories labeling." *Public health nutrition* 11:203-213.
Wansink, B., and P. Chandon. 2006. "Can “low-fat” nutrition labels lead to obesity?" *Journal of Marketing Research* 43:605-617.
Yen, S.T. 2005. "A multivariate sample-selection model: estimating cigarette and alcohol demands with zero observations." *American Journal of Agricultural Economics* 87:453-466.
Zhen, C., and X. Zheng (2015) "The effects of expert opinion on consumer demand for goods with credence attributes: Evidence from a natural experiment." In *Working Paper, Department of Agricultural Economics, North Carolina State University*

Working Paper, Department of Agricultural Economics, North Carolina State University.

Notes

[i] While more than 1,500 U.S. supermarkets assign Guiding Stars to food products.

[ii] The analysis by Nikolova and Inman (2015) may have some econometric limitations: (1) their analysis may generate biased parameter estimates because NuVal score is omitted in their volume model, (2) unobserved product and household characteristics might be related with the error term, and (3) the score model might suffer of endogeneity problems because the score information is included in both sides of the regression. In reality, score might not be a choice variable.

[iii] For the "card" approach to work well, the scanner data company has to have agreement with most retailers in the market for them to supply purchase data. Therefore, this approach works better in small markets where there are a limited number of retailers.

[iv] Although another inputs such as housing, diet, recreation, smoking, medical care, and alcohol consumption influence health level Grossman, M. 1972. "On the concept of health capital and the demand for health." *Journal of Political economy* 80:223-255.; we treat diet as the most important market good in the gross investment function to evaluate the role of diet on health.

[v] The time spent cooking at home instead of buying processed foods is captured by *TH*.

[vi] Although price endogeneity is a minor problem for a household-level analysis because price is exogenous to consumers (Zhu, Lopez, and Liu, 2016), we include UPC fixed effects in the model to rule out the case that unobserved product characteristics are correlated with price.

[vii] The introduction of NuVal labeling may affect food processors and retailers' responses via product reformulation and marketing and sales strategies, respectively, (See Berryman, P. 2014. *Advances in Food and Beverage Labelling: Information and Regulations*: Elsevier. Therefore, we control for advertising and price discounts (See ibid).

[viii] A two-part model is an analog to hurdle-models for zero-inflated count data.

[ix] We estimate the log-gamma distribution because fits conditional purchases (measured in volume) better than the log-normal distribution.

[x] The treatment store is owned by a regional grocery chain. Among the control stores, two are owned by a local food Co-op, one by another regional grocery chain, and two each by a local independent owner.

[xi] Our data of cold cereal products have a combined market share of approximately80%.

[xii] We used UPCs that exist before and after post label period so we are able to control for the possibility that post label period the store might have a higher introduction of healthier products (i.e., private brands).

[xiii] We select the cutoff point of 50 for two reasons. First, the NuVal scores are designated to range from 1 to 100; therefore, it is reasonable to think that consumers may believe that UPCs with scores equal or higher than 50 are healthy ones and those below 50 are unhealthy. Second, scores for cold cereal in our sample range from 10 to 91, therefore 50 is a natural cutoff point.

[xiv] Using 185% of federal poverty guidelines as a cutoff, we classify shoppers as low-income or high-income households.

[xv] The scale is different from the estimated scale parameter (0.6) for a log-normal variable obtained using the mean (605) and the standard deviation (435) of purchases.

[xvi] There are 6 groups of households out of 16 whose sample size is less than 1%.