SHELF NUTRITION LABELS AND CONSUMER BEHAVIOUR

by

GRACE MELO

(Under the Direction of Chen Zhen and Greg Colson)

ABSTRACT

The objective of this dissertation is to examine the effect of the NuVal shelf nutrition label—a food scoring system that presents a parsimonious summary of the nutrition profile of food products—on consumer purchasing behavior: quantity and participation decisions across different household groups. Chapter one estimates the impact of the label on breakfast cereal purchases by using a Two-Part Model. Chapter two explores the effect of the label on frozen dinner purchases by estimating a Two-Part Model. This chapter also explores the effects of the label across the purchase intensity employing quantile regression analysis.

INDEX WORDS: NuVal, Consumer Choice, Shelf Nutrition Labels, Two-Part Model, Quantile Regression, Breakfast Cereal, Frozen Dinner, Preference Heterogeneity

SHELF NUTRITION LABELS AND CONSUMER BEHAVIOUR

by

GRACE MELO

B.S., Zamorano University, Honduras, 2010

M.S., University of Georgia, 2013

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2017

SHELF NUTRITION LABELS AND CONSUMER BEHAVIOUR

by

GRACE MELO

Major Professor:	Chen Zhen
	Greg Colson
Committee:	Travis Smith
	Joshua Berning

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

DEDICATION

Dedicated to God, my lovely family, and my friends who have supported me during this academic experience and to my professors in Zamorano University and The University of Georgia who have challenged me to go beyond what I thought I could do.

ACKNOWLEDGMENTS

Furthermore, I want to recognize all my professors in the Department of Agricultural and Applied Economics and in other departments who have planted passion in me about Economic sciences. This was a crucial factor that influenced my decision to apply to Ph.D. Thanks to Dr. Colson, Dr. Zhen, and Dr. Ramirez for sharing not only your knowledge but also giving me advice. Thanks to all who believed in me and made it possible to achieve my goal and develop my educational and professional skills.

Very special thanks to my family for supporting me in all possible ways. Thanks to God for wisdom and strength during this challenging and at the same time stimulating path.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS v
LIST OF TABLES viii
LIST OF FIGURES ix
CHAPTERS

1. THE IMPACT OF NUVAL SHELF NUTRITION LABELS ON FOOD CHOICES: EVIDENCE FROM COLD CEREAL PURCHASES

    Introduction 1
    Shelf Nutrition Labeling and Consumer Choice 3
    Store- and Household-Level Scanner Data 6
    No-Purchase Option in Scanner Data 7
    Theoretical Framework 8
    Empirical Framework 14
    Data 18
    Empirical Results 23
    Conclusions and Implications 31

2. THE IMPACT OF NUVAL SHELF NUTRITION LABELS ON FOOD CHOICES: EVIDENCE FROM FROZEN DINNER PURCHASES

    Introduction 40
    Shelf Nutrition Labeling and Processed Foods 43
    Theoretical Framework 44
    Empirical Framework 52
    Data 59

Empirical Results 63

Conclusions and Implications 75

REFERENCES 89

APPENDIX 100

A Log-Gamma Distribution for the second part of the TPM 100

B NuVal Effects Log-Gamma Distribution 101

C Unconditional Own-Price Elasticity 101

D Log-Normal Distribution for the second part of the TPM 102

E NuVal Effects Log-Normal Distribution 102

F Conditional Own-Price Elasticity 103

G Unconditional Own-Price Elasticity 104

H Partial Effects of the TPM 105

LIST OF TABLES

Page

Table 1. 1 Difference between Non-NuVal and NuVal Stores 33

Table 1. 2 Summary Statistics for the Regression Analysis 34

Table 1. 3 Comparison of Estimation Results 35

Table 1. 4 TPM Estimation Results with Heterogeneous Effects 36

Table 1. 5 Conditional NuVal Effects on Purchases across Demographic Groups 37

Table 1. 6 Unconditional NuVal Effects on Purchases across Demographic Groups 37

Table 2. 1 Difference between Non-NuVal and NuVal Stores 77

Table 2. 2 Summary Statistics for the Regression Analysis 78

Table 2. 3 Estimation Results 79

Table 2. 4 Purchasing Patterns across Households 81

Table 2. 5 Conditional NuVal Effects on Purchases across Demographic Groups 82

Table 2. 6 Unconditional NuVal Effects on Purchases across Demographic Groups 82

Table 2. 7 Quantile Regression Results across Demographic Groups 83

LIST OF FIGURES

Page

Figure 1. 1 Example of a Price Tag with NuVal Score for Cold Cereal at a NuVal store 38

Figure 1. 2 The Marginal Efficiency of Health Capital (MEC) Demand Curve 38

Figure 1. 3 Distribution of Weekly Purchases of Cold Cereal during the Sample Period............ 39

Figure 1. 4 Distribution of NuVal Scores of the UPCs in the Sample of Cold Cereal Products.. 39

Figure 2. 1 Utility Maximization and Opportunity Cost.. 84

Figure 2. 2 Distribution of Weekly Purchases 84

Figure 2. 3 Distribution of NuVal Scores of the UPCs in the Sample of Frozen Dinner Products 85

Figure 2. 4 Distribution of Conditional Purchases across Household Groups: no college households (no_college), single-person households (single), families with children (children) and low-income shoppers (low_inc)... 86

Figure 2. 5 Quantile Regression Results for Price 87

Figure 2. 6 Quantile Regression Results for the Effects of NuVal Scores 88

CHAPTER 1

THE IMPACT OF NUVAL SHELF NUTRITION LABELS ON FOOD CHOICES: EVIDENCE FROM COLD CEREAL PURCHASES

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 like nutritional quality. The Nutritional Facts Label (NFL)—a government initiative to help U.S. consumers to make healthy food choices —has signaled the nutritional quality of foods since 1994. Nevertheless, more than half of U.S. consumers do not read the NFL (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 consumers' use 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 considered the development of a standardized nutrition label that provides clear and concise nutritional information on these foods (FDA, 2009).

Summary shelf nutritional labels summarize the NFL information in a simple manner; therefore, compared to current labeling systems that are nutrient-specific (e.g., FOP), these labels might be a better alternative to informing consumers of the nutritional value of foods. Evidence regarding whether NuVal, a summary shelf label system, is effective in influencing consumer behavior when visiting the grocery 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 determines whether 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). Understanding the effect of a summary labeling system across different household groups is important when considering a nutritional labeling policy that aims to improve consumers' understanding of nutritional information. More critical for t for policy making is evaluating whether those households that have a poor understanding of the nutritional information provided in current labels (e.g., low-income households with poor educational attainment) improve their food choices with the use of summary labels.

In addition, because it is possible that NuVal labels can make households more likely to purchase healthier products their analysis based on data contingent on households making purchases fails to capture the overall impact of the NuVal labels. In this article, we present unique evidence assessing the impact of NuVal shelf nutritional labels on consumer behavior across different household groups. In lieu of using a 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 impact of NuVal labels on the food choices across different household segments.

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 methods to analyze scanner data. Then, we provide a theoretical framework of the demand for health that served as the basis to explain asymmetric label effects across heterogenous households in our empirical model. Next, we summarize the data features 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 products (Hersey et al., 2013). In addition, consumers may associate descriptors such as *low-fat* with a poor taste (Berning et al., 2010, Teisl et al., 2001, Wansink and Chandon, 2006) 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 (Berry et al., 2015, 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. Although nearly 2,000 supermarkets have adopted NuVal, most studies focus on the Guiding Stars label, which is the second most used shelf nutrition labeling system in U.S. stores after NuVal (Anand, 2016). Guiding Stars, currently adopted in more than 1,500 supermarkets, 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 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 factor (Katz et al., 2010, NuVal, 2012). Like 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 indicates 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 across different demographic groups. The ability to identify heterogeneous NuVal effects across different households 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) 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 features that may correlate with the explanatory variables (e.g., price). In addition, their study based on seemingly unrelated regression estimation has additional econometric limitations: (1) observed and unobserved household characteristics driving purchases are not captured in their system, and (2) the NuVal score information is included on both sides of their nutritional content regression, which raises endogeneity problems and inconsistent parameters.

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; however, this does not eliminate the problem (Gupta et al., 1996). In fact, for the "card" approach to work well, the scanner data company should have an agreement with most retailers in the market for them to supply purchase data.

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. Luckily, 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.

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 several 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, *X* represents the inputs to produce Z. The other inputs TH and T are time inputs, and K is the stock of capital. Although other inputs such as housing, diet, recreation, smoking, medical care, and alcohol consumption also influence health level, we treat diet as the most important market good in the gross investment function to evaluate the role of diet on health.

Because production functions are homogeneous of degree 1 in the inputs, the production function of gross investment in health can be defined as:

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

where $t = \frac{TH}{D}$ and the marginal products of the inputs are $\frac{\partial I}{\partial TH} = \frac{\partial g}{\partial t}$ and $\frac{\partial I}{\partial D} = g - \frac{\partial g}{\partial t} t$

The budget and time constraints can be defined as:

(6) $P_0 D_0 + V_0 X_0 + \frac{P_1 D_1 + V_1 X_1}{1+r} = TW + A$

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

where *P* and *V* are the input prices of *D* 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, *TL* is the time lost due to illness, and *T* is number of non-working hours remaining to produce *Z*. We assume that $\frac{\partial TL}{\partial H} < 0$ and $\frac{\partial I}{\partial TH} > 0$. The time spent in improving health (e.g., preparing and/or eating a nutritious diet) is captured by *TH*.

We can define the full wealth constraint as:

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

Substituting equations (6) and (7) into equation (8), we can define the full wealth constraint as:

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

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:

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

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

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

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

Therefore, equation (11) can be written as:

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

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 (12) 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:

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

The FOCs for gross investment are:

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

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

where $\frac{\partial M}{\partial I} = \frac{1}{g - \frac{\partial g}{\partial t}t}$. Substituting for λ from equation (15) into equation (14) we obtain:

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

Equation (15) 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 (12) as:

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

Equation (17) can be arranged as follows:

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

Equation (18) 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 the approach by Grossman (1972) and ignore the consumption of health from now on (i.e., Pure Investment model). Then, equation (18) and the full wealth constraint can be reduced to equations (19) and (20):

(19) $$\gamma_i = \frac{G_i W_i}{\pi_{i-1}} = r - \frac{\pi_i - \pi_{i-1}}{\pi_{i-1}} + \delta_i \frac{\pi_i}{\pi_{i-1}}$$

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

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, the greater the person's value of a rise 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 1.2). 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:

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

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

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:

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

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 1.2). Consequently, 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 to produce gross capital investment will increase.

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

We select breakfast cereal to estimate the impact of NuVal labels because of 1) its important contribution to the nutritional quality of the diets of children and adolescents in the U.S. (Morgan et al., 1986), 2) the WHO guidelines recommended a reduction of daily sugar intake to 5% of total calories to address obesity and health-related conditions (World Health Organization, 2015). Cold cereal is one of the top sources of added sugars for younger children

(Reedy and Krebs-Smith, 2010), and 3) its high purchase volume (market share of approximately 80% in our data) and large variation in NuVal scores (min 10, max 91).

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. We model households' purchases using a TPM, a two-part analysis. We also estimate purchases using one-level analyses: Conditional One-level Analysis and Unconditional One-level Analysis to compare the advantage of employing a TPM over conventional one-level analysis. Finally, we also estimate stores sales to identify the impact of NuVal scores on total sales of the store that adopted NuVal.

Conditional One-level Analysis

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

$$(24) \qquad v_{hitr} = a_i + a_t + a_r + a_h + b_1\, P_{itr} + b_2\, Adopt_{itr} + b_3\, Adopt_{itr} * Score_i + M'_{itr}\gamma + \epsilon_{hitr}$$

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 volume of UPC *i* at store *r* in week *t*; $Adopt_{itr}$ is a dummy equal to one if store *r* had posted the 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 NuVal effect or salience effect of the NuVal label on the labeled UPCs at the NuVal store. This salience effect only indicates to consumers that the product has received a NuVal score; hence it does not capture the NuVal 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.

Studies have shown that FOP labels have influenced food processors and retailers' decisions related to product reformulation and marketing and sales strategies (Berryman, 2014). Therefore, one might be concerned that the non-NuVal stores and the NuVal store could have increased marketing activities after NuVal labels were implemented; therefore, we include the vector M_{itr} that controls for advertising and price discounts in store r for UPC i in week t.

Unconditional One-level Analysis (Naïve)

One approach to accounting for no-purchase outcomes in a simple manner is regressing the choice variable y_{hitr}, which includes both non-zero purchases and zero purchases, on the set of regressors in equation (24). 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 (24) using store-level data. This model, which is conditional on observing stores sales, is specified as:

$$V_{itr} = a_i + a_t + a_r + b_1 P_{itr} + b_2 Adopt_{itr} + b_3 Adopt_{itr} * Score_i + M'_{itr}\gamma + \epsilon_{itr} \quad (25)$$

where V_{itr} is sales volume at store *r* in week *t* of UPC *i* and the remaining variables are defined as in equation (24).

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 (24). 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, an analog to hurdle-models for zero-inflated count data, 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 1.2) followed by right-skewed data (Figure 1.3). This non-normal distribution of the non-zero purchase data can be accommodated in the conditional part of the TPM.

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

$$(26) \qquad y_{it}^{*} = f(X_{it}, \beta) + u_{2i} + \epsilon_{it}$$

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

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

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 (26) 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 as:

$$\text{(27)} \qquad \text{logit}\big(\text{prob}(d_{it} = 1)\big) = logit(p_{it}) = \log\left(\frac{\pi_{it}}{1-\pi_{it}}\right) = Z_{it}\alpha + u_{1i} + \epsilon_{it}$$

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 (24).

$$\text{(28)} \qquad \text{E}(Y_{it}|d_{it} = 1\,) = q_{it} = f(X_{it}\beta) + u_{2i} + \ \epsilon_{it},\, q_{it} \sim N(\mu, \phi\,)$$

where f is a monotone increasing function (e.g., log-normal or log-gamma) that will make μ approximately Gaussian (i.e., normal) and ϕ is a dispersion parameter. We employ the log-gamma distribution because fits conditional purchases (measured in volume) for cold cereal better than the log-normal distribution.

Heterogeneous Consumers' Responses to NuVal

Health concerns and nutrition knowledge are some of the predictors of label use (Drichoutis et al., 2006).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 (24) are incorporated as follows:

(29) $$v_{hitr} = a_i + a_t + a_r + D_h' b_0 + b_1 P_{itr} + (b_2 + D_h' b_4) Adopt_{itr} + (b_3 + D_h' b_5) Adopt_{itr} * score + M_{itr}' \gamma + \epsilon_{hitr}$$

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 (29) 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). 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 (NuVal store) and no other stores in the city adopted either Guiding Stars or NuVal labels during our sample period. The NuVal store is owned by a regional grocery chain. Among the non-NuVal stores, two are owned by a local food Co-op, one by another regional grocery chain, and two each by a local independent owner.

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 NuVal 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 non-NuVal period in our analysis. Table 1.2 indicates that our sample consists of 6 grocery stores, 2652 households, and 186 UPCs that were sold during both NuVal and non-NuVal periods. Specifically, we used UPCs that exist before and after post label period so we can control for the possibility that post label period the store might have a higher introduction of healthier products (i.e., private brands). The NuVal scores for cold cereal in our sample range from 10 to 91, with an average of 30.

Differences between NuVal and Non-NuVal Stores

Before conducting the regression analysis, we first use summary statistics to compare changes in purchases and sales after NuVal adoption without control for covariates. Table 1.1 provides the average weekly quantity purchased at the NuVal and non-NuVal 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 NuVal and non-NuVal stores during the non-NuVal and NuVal periods. The first two columns of the table report the mean purchases and sales for all labeled UPCs, regardless of the NuVal scores for the household- and store-level data. In the remaining columns of Table 1.1, we investigate whether the NuVal effect is different for UPCs with higher and lower NuVal scores.

First, to examine the differences between the non-NuVal stores and the NuVal store, column 1 provides the summary statistics for the cold cereal products purchased and sold in these stores during the non-NuVal period. The summary statistics for household-level data shows that households purchased more volume of cold cereal per week in the non-NuVal stores than in the NuVal store during the non-NuVal period (619 vs. 592 grams). However, the summary statics of store-level data indicates that every week the NuVal store sold more volume of every UPC than the non-NuVal stores during the non-NuVal period (9440 vs. 5768 grams). During the study period, the NuVal store carried 96% of labeled UPCs of our sample while the non-NuVal stores carried 99%. Overall, this indicates that although the NuVal store did not carry all UPCs in the sample, it was the market leader for cold cereal in terms of weekly sales in the study town during the 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 1.1 show the price per gram of product, which is the price paid by the household, 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 NuVal and non-NuVal 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 NuVal and the non-NuVal stores during the non-NuVal period were $0.96 and $0.91, respectively. The mean scores of the UPCs in the NuVal store and the non-NuVal stores were about 29 (i.e., 29.31 and 29.58, respectively). In terms of marketing activities, the non-NuVal stores made more promotional efforts than the NuVal store. The non-NuVal stores offered more discounts of at least 5% (25% vs. 14%) and advertised more (11% vs. 2%) cereal products than the NuVal store during the non-NuVal period. The differences in means between the non-NuVal and NuVal stores, in terms of marketing efforts, indicate the need to control for these covariates using a regression.

Differences in Means

We compare purchases between non-NuVal and NuVal periods at the NuVal 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. We select this cutoff for the following 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.

As expected, 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 1.4), the overall effect was a slight increase in households' purchases by 12 grams at the NuVal store after the NuVal label adoption.

Comparing sales between non-NuVal and NuVal periods at the NuVal store, we observe that the NuVal 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 labeled UPCs at the NuVal store after the label was adopted.

During the same sample period, in the non-NuVal stores, the average volume sold for each labeled 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 the change in means, suggests that posting NuVal labels increased consumer demand for the labeled UPCs by 52.38 grams, or about 9% of 592 grams, the average volume purchased in the NuVal 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 non-NuVal and NuVal stores (i.e., -\$0.04 and -\$0.07 per 100 grams of product, respectively). Similarly, promotional activities decreased at both the NuVal and non-NuVal 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 non-NuVal stores and by 2 percentage points at the NuVal store.

A comparison between store- and household-level data show 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 NuVal store compared with the non-NuVal period (3% vs. 9% of sales and purchase volume, respectively). Second, the store data indicates that the NuVal 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 non-NuVal stores than in the NuVal 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 purchases. We use regression analysis to control for these factors and estimate the NuVal label effect.

Sample Characteristics

Table 1.2 reports the summary statistics for all variables used in the analysis for cold cereal. The summary statistics table 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% had 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, 19% have children, 57% have not attended college, 7% of households have heads with smoking habits, and the average household consists of two members. We classify shoppers as low-income or high-income household using 185% federal poverty guidelines as a cutoff.

Empirical Results

One-level Analysis

Table 1.3 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 1.3 also reports the estimated NuVal 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 NuVal effect becomes positive. The negative sign of the parameter estimates for *Adopt* in columns 1, 2, and 4 indicate that the estimated NuVal 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 NuVal 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 NuVal effect for a UPC at the average NuVal score is negative (-165.37). This indicates that compared with sales in the non-NuVal stores, sales for a UPC with a score of 30 (i.e., the mean score) decrease by 165 grams in the NuVal store with the posting of the NuVal score. Zhen and Zheng (2015) reported an increase in yogurt sales for NuVal 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 purchases for all NuVal scores. Therefore, the estimated NuVal effect at the mean score is positive (44.67). This indicates that compared with purchases in the non-NuVal 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 NuVal effect is positive (0.08).

To be able to compare the NuVal effects across these three models (store-level data, unconditional, and conditional models), we compare the estimated NuVal effects with the effect of a price change. The estimated NuVal 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 NuVal 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 NuVal 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 findings support the results of Table 1.1 that also shows a larger effect (measured by the difference in means) of the NuVal labels when using household-level data rather than store-level data.

As discussed, the unconditional model in our study introduces bias if the purchasing decision and the decision on 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 1.3. The scale parameter in the quantity decision equation is 0.14 which is statistically different from 1; therefore, the hypothesis of an exponential distribution for the data is rejected. In addition, the scale value (0.14) is different from the scale parameter if a log-normal distribution is assumed (i.e., 0.6), which was calculated using the mean (605) and the standard deviation (435) of purchases.

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 labeled 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 NuVal 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 because of the NuVal scores, we estimate the unconditional NuVal effect. Using the TPM parameter estimates, we estimate the NuVal effect (conditional and unconditional) for a UPC that does not have a price mark-down or a coupon (i.e., *PR* and *Ad* are zero) at the average values of score and price (i.e., 30 and $0.01 per gram of product). Calculations of the conditional and unconditional NuVal effects for a log-gamma distributed variable are shown in Appendix B. The estimated conditional and unconditional NuVal effects are 0.001 and 0.0001 grams, which are equivalent to a 6% and 8% increase with respect to the estimated purchase volume before NuVal adoption. Although the unconditional NuVal effect is lower compared with the estimated NuVal effects obtained in the unconditional and conditional models (columns 2 and 3), the impact of the NuVal label expressed in percentage change is not trivial, especially after considering changes in the purchasing probabilities.

Altogether, the TPM results indicate that NuVal labels not only affect the quantity decision, but they also affect the purchasing intention. Hence, the approach to explain consumer behavior with model conditional on purchases will only reflect the partial impact of these 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 the 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 1.4 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 1.4 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 of 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 NuVal 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 1.2. The results for the heterogeneous effects of the participation equation (column 3) for this household subsample shows that among low-income households, families without any college education and household with 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 a higher volume of healthier products compared with high-income families with no children and larger households when NuVal scores are posted.

Conditional NuVal Effects on Purchases across Demographic Groups

Table 1.5 reports the estimated conditional NuVal effects on purchases across household groups. We report the NuVal effects for those household groups with a sample size greater than 1% only. In our sample, there are 6 groups of households out of 16 whose sub-sample size is less than 1%.

Column 1 indicates the proportion of each household group in the sample. Columns 2 to 4 indicate the NuVal effects (expressed in percentage volume change with respect to the predicted purchase volume before NuVal adoption) for UPCs with different NuVal scores levels (min 10, mean 30, max 91).

The conditional analysis indicates that the NuVal effects for a UPC with the minimum score value of 10 are negative for all shoppers. The negative sign of the NuVal 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., UPC with a NuVal score of 10) by 36% due to the NuVal labels. They also experience the largest conditional NuVal effect when a UPC is assigned the highest NuVal score (i.e., 91). They increased purchases of this UPC by about 59% after the NuVal adoption. The second largest improvement in 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 NuVal Effects on Purchases across Demographic Groups

While the conditional NuVal effects are informative, they do not account for changes in shoppers' likelihood to purchase UPCs with higher and lower NuVal scores. The estimated unconditional NuVal effects reported in Table 1.6 indicate to some extent different results compared with the conditional NuVal effects in Table 1.5. Differences in the estimated NuVal effects between the conditional and unconditional analysis are caused because the parameter estimates for the heterogeneous NuVal 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 1.4).

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 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 NuVal store in this small Midwestern town.

Conclusions and Implications

With the intent of the FDA to redefine the nutritional claim “healthy” and develop a standardized, science-based criteria symbol that provides clear and concise nutritional information of processed foods, examining the effectiveness of summary nutrition labeling systems on improving food choices is critical. Employing purchase data related to a supermarket’s voluntary adoption of NuVal— a 1 to 100 numeric summary shelf label system, we estimate a Two-Part Model (TPM) to identify the effect of the NuVal label on consumer purchasing decisions for cold cereal. In addition, we test whether the households’ purchases are representative of sales at the stores in our sample by estimating 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 NuVal effects between store- and household-level data analyses were to some extent different. First, we found differences in the NuVal effects between sales and purchases using summary statistics. 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 asymmetric NuVal effects 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 outcome might be because the gap of nutritional information *ex-ante* of lower-income households is larger compared with other household groups. Our findings suggest that providing interpretative summary nutrition information on the overall nutritional value of foods can be an effective way to improve consumer choices.

Tables

Table 1. 1 Difference between Non-NuVal and NuVal Stores

		Volume			Volume & Score<50 (176 UPCs)			Volume & Score>=50 (10 UPCs)			Price (100 grams)		Ad		PR		Score
*Period		Non-NuVal	Trt.	Change	Non-NuVal	Trt.	Change	Non-NuVal	Trt.	Change	Non-NuVal	Trt.	Non-NuVal	Trt.	Non-NuVal	Trt.	
Household-Level Data																	
Non-NuVal 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
NuVal 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%							
Store-level Data																	
Non-NuVal 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
NuVal 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 NuVal, and Diff. denotes Difference. * Because UPCs were assigned scores on different dates, UPCs scored after the NuVal adoption have unique NuVal and non-NuVal periods.

Table 1. 2 Summary Statistics for the Regression Analysis

Variables	Description	Mean	S.D.
Dependent Variables			
V	Sales volume, number of equivalised 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 equivalised units purchased of UPC *i* in time *t* from retailer *r* by household *h*	605.37	435.08
Explanatory Variables			
NuVal Variables			
Adopt	Adopt=1 if UPC *i* had been assigned a score during the NuVal period at the NuVal 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 (grams)	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 both heads of a household have not attended college, 0 otherwise	0.57	0.50
Smoke	Smoke=1 if both heads of a household 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 1. 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
NuVal 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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. [a] Unconditional NuVal Effects

Table 1. 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 1. 5 Conditional NuVal Effects on Purchases across Demographic Groups

Household Group	% Sample	NuVal 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: NuVal 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 1. 6 Unconditional NuVal Effects on Purchases across Demographic Groups

Household Group	% Sample	NuVal 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: NuVal 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. 1 Example of a Price Tag with NuVal Score for Cold Cereal at a NuVal store

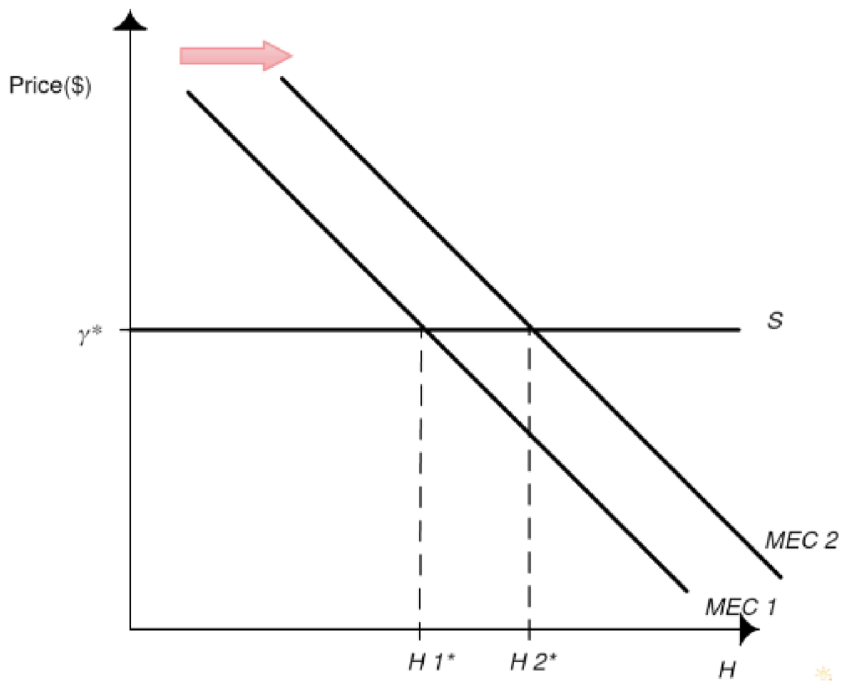

Figure 1. 2 The Marginal Efficiency of Health Capital (MEC) Demand Curve

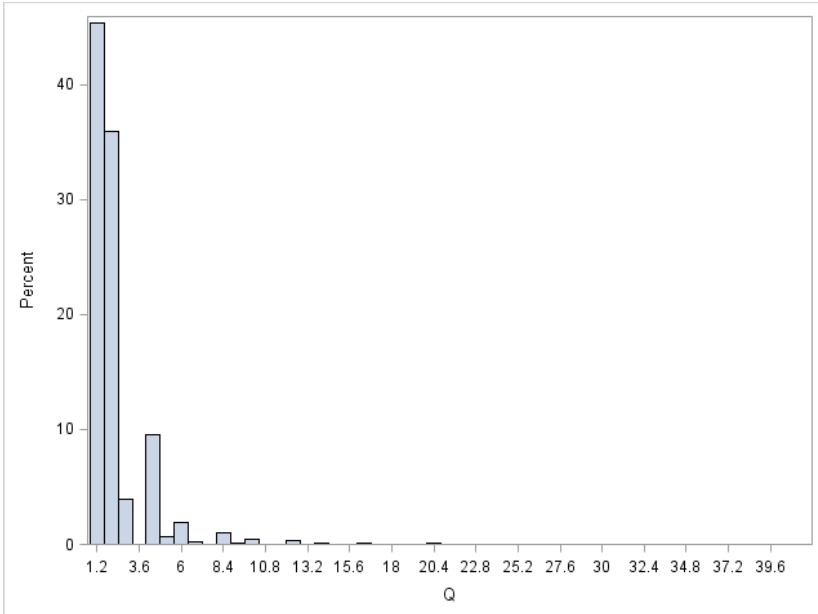

Figure 1. 3 Distribution of Weekly Purchases of Cold Cereal during the Sample Period

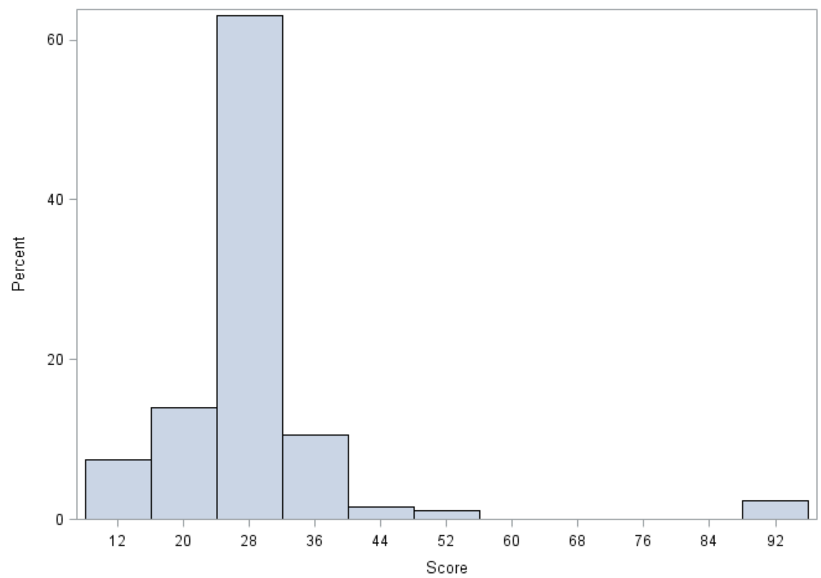

Figure 1. 4 Distribution of NuVal Scores of the UPCs in the Sample of Cold Cereal Products

CHAPTER 2

THE IMPACT OF NUVAL SHELF NUTRITION LABELS ON FOOD CHOICES: EVIDENCE FROM FROZEN DINNER PURCHASES

Introduction

The increase in women's labor force participation has amplified the opportunity cost of time and reduced food production at home (Aguiar and Hurst, 2006). This trend is reflected by the popularity of convenience foods such as ready-to-eat and take-out meals in recent years (Bowers, 2000). Unfortunately, this inclination for convenience, easy-to-prepare, and away-from-home foods has been related to a poor diet quality (Todd et al., 2010).

At the same time, the increasing evidence of a diet-health relationship has been an important factor influencing consumers' preferences shifts towards healthier food options (Gagliardi, 2015). The food industry has responded to these changes by developing "healthier" processed products, which has created a market of processed foods with a wide-ranging nutritional level (Gagliardi, 2015). To be able to differentiate these "healthier" foods, U.S. food manufacturers and retailers have developed voluntary label systems such as Front-of-Package (FOP) labels that supplement the required Nutrition Facts labels. However, differences across the variety of voluntary labeling systems and nutritional and health claims have made it difficult for consumers to distinguish the nutritional quality of foods. To facilitate consumers' understanding of the nutritional value of foods, the Institute of Medicine has called for a single, standardized, and universal FOP symbol (IOM, 2011). Based on this recommendation, the WHO advised a more unified approach to FOP labeling and the FDA considered the possibility of developing a mandatory standardized symbol that provides consumers with information to enable them to easily and quickly make food choices consistent with health recommendations (FDA, 2009, Scott-Thomas, 2014).

A recent approach to communicating nutrition information in a simple and concise manner is summary shelf nutrition labels. Consumers, on average, make more than 200 food-related decisions (Wansink and Sobal, 2007), therefore, summary shelf labels may ease households' decision-making process when buying processed foods, especially for time-constraint households, who tend to overlook existing nutritional labels (Grunert and Wills, 2007). The summary nutrition scores on the shelf labels are algorithm-based ranking systems that use Nutrition Facts label information, nutrition information not on the Nutrition Facts label, and diet-disease evidence to rank products by their nutritional value (Armstrong, 2010). In the U.S., Guiding Stars and NuVal are the main shelf labels, which have been displayed on the shelf price tag at participating stores since 2006 and 2008, respectively. Because the presence of these label systems is relatively new compared to other nutrition labels and nutrient claims, research examining the effect of these labeling systems on actual consumers' choices is scarce. Moreover, previous empirical work has two main limitations: 1) zero purchases are ignored and 2) the impact of the label is assumed to be homogenous across households and levels of purchase intensity. In this study, we test the ability of NuVal scoring systems to improve households' frozen dinner purchase decisions using a Two-Part Model (TPM) and quantile regressions. These estimation methods based on the analysis of household-level data allow us to address the two main empirical limitations of previous work.

Early studies that test the ability of shelf nutrition labels in improving food choices are limited to Guiding Stars (Rahkovsky et al., 2013, Sutherland et al., 2010), which is a 4-point rating system with 0 and 4 stars indicating least healthy and healthiest products, respectively. In contrast, NuVal uses a 100-point scale, which may allow consumers to better differentiate the nutritional quality of products. To date, there are only two empirical studies that evaluate the impact of NuVal labels on food choices. One study is based on store-level sales of yogurt (Zhen and Zheng, 2015), and the other study estimates households purchases for eight food categories ignoring corner solutions and preference heterogeneity in the demand for these foods (Nikolova and Inman, 2015). Neither of these studies allows for heterogeneous label effects due to differences in levels of purchasing intensity and demographics across households. Understanding heterogeneity of households' responses to NuVal labels is important for a potential targeted approach to nutrition labeling and education.

In our study, we estimate the impact of the NuVal labels on household purchases of frozen dinner and make the following contributions. First, we estimate the impact of the labels employing a TPM that allows us to analyze corner solutions due to zero purchases and capture the impact of the NuVal labels on households' purchase and quantity decisions. Second, we estimate the impact of the labels across different demographic groups. We specifically test the ability of NuVal to generate larger improvements in food choices of consumers with poor knowledge of nutritional information relative to consumers with a better understanding of this information (i.e., low-income and low-educated households vs. high-income and educated households). We test whether time-constrained household benefit more from NuVal scores (i.e., full-time employees vs. part-time employees vs. retired households). Finally, we test whether the level of preferences level for frozen dinner purchases measured by volume influence the effect of NuVal (e.g., light users vs. heavy users). To test different responses between heavy users and light users, we estimate quantile regressions for each socio-demographic group.

The remainder of the paper is organized as follows. We review the literature on the effect of shelf nutritional labels on processed foods. Next, we provide a theory on time allocation and relate this theory to the demand for convenience foods such as frozen meals. This is followed by a description of the empirical model and scanner data. Then, we present and discuss the empirical results. The final section concludes the study.

Shelf Nutrition Labeling and Processed Foods

The food industry has responded to consumers' preferences for healthier products by developing products that have higher nutritional quality and increasing the visibility of FDA-approved health and nutrition claims in product packaging (e.g., Front-of-Package labels). As a result, the market for processed foods is saturated with goods of a wide range of nutritional quality (Gagliardi, 2015). However, labeling information of processed foods has been subjected to controversy related to their transparency (Grunert and Wills, 2007). Subsequently, the FDA considered the option of developing a standardized mandatory symbol that provides consumers with concise information to enable them to easily and quickly make food choices consistent with health recommendations (FDA, 2009). Moreover, given publicly expressed concerns about voluntary labeling systems providing misleading information, the FDA has been seeking to redefine the nutrient content claim "healthy" (FDA, 2016). However, redefining this claim raises controversy about whether products with high levels of added sugar or that contain saturated fat should be allowed to be associated with this claim (Watson, 2015, Watson, 2017).

Nutritional shelf label systems are voluntary labels that provide concise and simplified nutritional information in the shelf tag along. The visibility of the shelf tags, containing also the price of the product, eases the transfer of information. Because they can reduce the costs of information and ease consumer cognitive process when shopping (Zhu et al., 2015), these labeling systems were recommended to promote healthy purchases among participants of food assistance in a USDA report (Gordon et al., 2014).

In the average American diet, calories from processed foods represent over half of total calories (Eicher-Miller et al., 2012), hence nutritional shelf labels might contribute to nutritional improvements, especially among households who heavily rely on processed foods to meet their caloric needs.

NuVal and Guiding Stars are two major shelf nutritional label systems implemented in U.S. grocery stores. These label systems are based on mathematical algorithms that rank products by their nutritional value. Most of the previous work on the effect of NuVal and Guiding Stars is based on sales data (Cawley et al., 2015, Rahkovsky et al., 2013, Sutherland et al., 2010, Zhen and Zheng, 2015). The main limitation of sales data analysis is that prevent us from testing the possibility that the labels have a different effect on each household group. The only study that, employed household-level data was conducted by Nikolova and Inman (2015); however, their study assumes homogenous responses and, more importantly, it does not address corner solutions due to zero purchases.

Theoretical Framework

Following the time allocation theoretical work by Becker (1965), we define a model in which households can produce basic commodities (e.g., preparing a meal) that affect their utility using the market goods (e.g., food ingredients) and time. The production functions of these commodities can be written as:

(1) $Z_i = f_i(x_i, t_i)$

where x_i are market goods and t_i are time inputs using in producing Z_i. In this model, a household is both producer and utility maximizer, and choose the optimal combination of input commodities to maximize utility:

(2) $U = U(Z_1, Z_2, \dots, Z_m)$

subject to the budget and time constraints:

(3) $\sum p_i x_i = T_w w + V = I$

(4) $\sum t_i = T - T_w = T_c$

where p_i is the price of x_i, T_w is the number of hours spent at work, w is the hourly wage rate, and V is the non-wage income. In the time constraint, T_c is the total time spent at consumption, and T is total time available. The production functions in equation (1) can be written as:

(5) $x_i = b_i Z_i$

(6) $t_i = a_i Z_i$

where a_i and b_i are the input of time and market goods per unit of Z_i. Substituting the constraints (4) to (6) into the budget constraint (3), we can define a general constraint as follows:

(7) $\sum p_i x_i + \sum t_i\, w = \sum (p_i b_i + a_i w) Z_i = V + Tw = S$

where $\sum (p_i b_i) Z_i$ represents expenditures on market goods and $\sum (a_i w) Z_i$ is the opportunity cost of spending time at consumption rather than at work. We can write equation (7) as:

(8) $\sum \pi_i\, Z_i = S$

where $\pi_i = p_i b_i + a_i w$ and $S = V + Tw$. The variable π_i represents the total price of unit Z_i and S is the *full income* achieved if all the time T were devoted at work.

Assuming that wage (w) is exogenous (i.e., uncorrelated with Z_i) and that a_i and b_i are fixed for given values of p_i and w, the equilibrium condition form maximizing the utility function in (2) subject to constraint (8) is given by:

(9) $\frac{\partial U_i}{\partial z_i} = \lambda \pi_i$

where λ is the marginal utility of income. The equilibrium condition in (9) assumes that a household spend all the time and resources to earn income with no time for consumption. A household with high-income, however, would give up work income to obtain additional utility from non-work activities (e.g., leisure). Therefore, the amount of income forgone, denoted as L, measures the cost of getting additional utility from these activities. The relationship between income forgone (*L)* and *full income* (*S)* can be denoted as follows:

(10) $L(Z_1, \dots, Z_m) = S - I(Z_1, \dots, Z_m)$

Substituting the budget constraint in (3) and equation (5) into equation (10), we can write:

(11) $\sum p_i b_i Z_i + L(Z_1, \dots, Z_m) = S$

If we assume a fixed average hourly wage, equation (11) can be reduced to equation (7). Therefore, with a constant average wage rate, we can use equation (6) and the condition $S = V + Tw$ to define income forgone (*L)* as:

(12) $L(Z_1, \dots, Z_m) = wT_c = w \sum a_i Z_i$

The equilibrium condition from maximizing the utility function subject to the constraint in (11) is given by:

(13) $\frac{\partial U_i}{\partial Z_i} = \lambda(p_i b_i + \frac{\partial L}{\partial Z_i})$

Considering that income can be allocated into market goods (x_i) and time (t_i), we can also define income forgone as $L = L(x_i, t_i)$. By taking the derivative of L_i with respect to Z_i, we obtain:

(14) $\frac{\partial L}{\partial Z_i} = \frac{\partial L}{\partial x_i}\frac{\partial x_i}{\partial Z_i} + \frac{\partial L}{\partial t_i}\frac{\partial t_i}{\partial Z_i} = \frac{\partial L}{\partial x_i} b_i + \frac{\partial L}{\partial t_i} a_i$

where $\frac{\partial L}{\partial x_i}$ and $\frac{\partial L}{\partial t_i}$ are the marginal forgone earnings from using an additional unit of a market good and time on producing z_i, respectively. Substituting equation (14) into equation (13) we obtain:

(15) $$\frac{\partial U}{\partial Z_i} = \lambda\left(p_i b_i + b_i \frac{\partial L}{\partial x_i} + a_i \frac{\partial L}{\partial t_i}\right) = \lambda\left(b_i(p_i + \frac{\partial L}{\partial x_i}) + a_i \frac{\partial L}{\partial t_i}\right)$$

where $b_i(p_i + \frac{\partial L}{\partial x_i})$ is the marginal cost of using market goods to produce Z_i and $a_i \frac{\partial L}{\partial t_i}$ corresponds to the marginal cost of using time for consumption of Z_i.

Figure 2.1 shows the equilibrium condition in (13) for a two-good world. In equilibrium, the slope of the *full income* curve, represented by equation (11) is equal to the slope of the indifference curve (i.e., $\frac{\frac{\partial U}{\partial z_1}}{\frac{\partial U}{\partial z_2}} = \frac{b_1(p_1+\frac{\partial L}{\partial x_1})+a_1\frac{\partial L}{\partial t_1}}{b_2(p_2+\frac{\partial L}{\partial 2})+a_2\frac{\partial L}{\partial t_2}}$).

Because we are interested in examining the effect of economic variables (e.g., income) of this theoretical model in the empirical section, we simplify the model by assuming that marginal forgone cost of using market goods is zero (i.e., $\frac{\partial L}{\partial x_i} = 0$) and that goods and time are used in fixed amounts (i.e., b_i and t_i are constants).

Hours of Consumption

This section evaluates the effect of changes in wages and market prices on the time spend on consumption (T_c) by examining the differences in the importance of forgone income among different commodities. The relative marginal importance of forgone income and time can be defined as:

(16) $$\alpha_i = \frac{\frac{\partial L}{\partial t_i} a_i}{p_i b_i + \frac{\partial L}{\partial t_i} a_i}$$

(17) $$\gamma_i = \frac{a_i}{p_i b_i + \frac{\partial L}{\partial t_i} a_i}$$

The marginal importance of forgone earnings (α_i) and time (γ_i) would be greater, the larger the number of hours used per unit of $Z_i(a_i)$; while they would be smaller, the larger the market price of goods and the number of goods used per unit of Z_i (i.e., p_i and b_i, respectively). Commodities with large α_i and γ_i can be considered *earnings-intensive* and *time-intensive* commodities. The commodity Z_1 in Figure 2.1 represents this type of commodities. Because α_i and γ_i tend to be positively correlated *time-intensive* goods are often *earnings-intensive* commodities only if $\frac{\partial L}{\partial t_i}$ and a_i (i.e., $\frac{\partial t_i}{\partial z_i}$) are positively correlated. For instance, forgone earnings would be more important for cooking dinner than buying dinner since the former uses more time per unit of dollar of goods than the latter.

Increase of wage

A percentage increase in wages by β would increase the cost per hour used in consumption by β for all commodities as it is demonstrated in equation (19):

(18) $L_0 = S - w - V = w' - w$

(19) $L_1 = (1 + \beta)(w' - w) = (1 + \beta)L_0$

Empirically, previous work has demonstrated that the value of household time and household income influenced food consumption away from home (McCracken and Brandt, 1987).

Increase of prices

An increase in wages will affect the relative prices of commodities as long as forgone earnings were not equally important for all goods. Specifically, the relative prices of commodities with relatively important forgone earnings, those with large α_i, would rise more. If Z_1 were the more *earnings-intensive* commodity, then households would consume less Z_1 and more Z_2, the less *earnings-intensive* commodity. In Figure 2.1 this change of allocation is represented by a shift from point A to B.

A shift away from *earnings-intensive* commodities and thus away from time-*intensive* commodities results in a reduction in the total time spent in consumption (T_c) and an increase in the time spent at work (T_w).

Summarizing, the importance of forgone earnings (α_i) are determined by the amount of time used per dollar of goods (a_i) and the cost per unit of time ($\frac{\partial L}{\partial t_i}$). These factors can vary among commodities and at different periods. For example, the cost of time for cooking would decrease when a household shifts from full-employment to retirement because of the decrease in earnings and thus declining the relative price of cooking. In contrast, the cost of time and the relative price of cooking would increase when a child is added to the family.

Prochaska and Schrimper (1973) tested Becker's model and found that as the value of a homemaker's time rose, measured by the wage rate constructed as a function of age and schooling, the household purchased more meals away-from-home and increased the demand for microwave, a time-saving durable item, and frozen dinner entrees.

Substitution between Time and Goods

Because substitution of inputs to produce z_i can take place, we relax the assumption that time and goods are used in fixed proportions in producing commodities (i.e., a_i and b_i).

For economic efficiency, the general equilibrium condition can be stated as follows:

(20) $$\frac{\frac{\partial U}{\partial x_i}}{\frac{\partial U}{\partial t_i}} = \frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial t_i}} = \frac{P}{\frac{\partial L}{\partial t_i}}$$

The equation in (20) indicates that the marginal rate of substitution ($\frac{\frac{\partial U}{\partial x_i}}{\frac{\partial U}{\partial t_i}}$) is equal to the marginal rate of technical substitution ($\frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial t_i}}$), which is equal to the ratio of the marginal costs of market goods to that of time. Hence, an increase in the cost of time will induce a shift towards market goods. For example, an increase in the value of a household's time may induce this household to spend less time cooking by using pre-cooked foods to prepare a dinner meal. If market goods are of *higher-quality*, then an increase in wages would increase the number of *higher-quality* goods purchased not only because of the substitution effect of time for goods but also because the effect of income on quality purchased.

One application of this theory is the substitution between frozen dinner and cooking. According to this approach, the cost of inputs to produce the commodity "*meal consumption at home*" is either the total expenditures of buying some foods to prepare a meal at the store and the forgone value of the time used to cook a meal or simply the price of a prepared meal (e.g., frozen meal), if the forgone value of the time used to prepare a frozen meal (e.g., heating) is close to zero. An increase in the price of frozen dinner relative to the cost to buy the ingredients, assuming that the value of time does not change, would reduce the cost of cooking dinner relatively to preparing a frozen meal, and thus a shift towards cooking would occur. A reduction in the value of time, assuming prices remain constant, would also shift production towards cooking.

Briefly, the importance of forgone earnings and the substitution between time and market goods may be quite relevant in analyzing consumption of frozen meals with different quality. The importance of foregone earnings also differs among households because of differences in income. Previous empirical work provides evidence that the increased opportunity cost of time as a result of a rise in income, rises expenditures of food-away-from-home and ready meals (Harris and Shiptsova, 2007, McCracken and Brandt, 1987).

Frozen Meal Shoppers

Energy-dense and nutrient-poor diets are associated with the consumption of processed foods and drinks (Todd et al., 2010), which are common foods in the diet of low-income shoppers (Drewnowski and Specter, 2004). Preferences for these foods can be related to the fact that the number of calories per dollar spent is higher for processed foods than for nutritious foods such as home meals. Alternately, time and effort are important barriers for preparing foods at home (Treiman et al., 1996). According to Devine et al. (2003) low-income households often have multiple jobs and tight job schedules, and therefore they might have higher preferences for convenience foods over home meals.

Because home-cooked foods can inexpensively be substituted for convenience meals (e.g., processed foods and fast-food meals), calories from these foods might be an important share of the diet of low-income and time-constrained shoppers. This could be linked to the possibility that time constraints may prevent participants of the Supplemental Nutrition Program for Women, Infants, and Children (WIC) and the Supplemental Nutrition Assistance Program (SNAP) from reaching healthy nutritional goals (Davis and You, 2011, Rose, 2007, Treiman et al., 1996).

To consider the fact that time constraints and the value of time influence consumer preferences for convenience foods, we conduct separate regression analysis of frozen dinner purchases for each household group in the next section. Household groups were classified based on their income level and employment status.

Previous work indicated that higher reading rates of labeling information were reported by shoppers when they bought a product for the first time (Grunert and Wills, 2007). This could be the case because information asymmetry on the characteristics of a product is greater among light-user shoppers compared to heavy-user households. Therefore, if light-user frozen dinner shoppers spend more time reading labels, we expect that these shoppers experience a smaller improvement of their food choices due to NuVal labels compared with high users. Therefore, in the empirical section, we also estimate quantile regressions for each household group, to test these heterogeneous effects across different levels of purchase intensity.

Empirical Framework

Employing a grocery retailer's voluntary adoption of NuVal shelf nutrition labels, we test whether posting summary nutrition scores on shelf labels improve consumers' purchasing choices of frozen dinner products. In the estimation of the TPM and quantile regressions, we control product, store, household, time unobservable characteristics, and product time trends that can affect purchases by including fixed effects at retailer, UPC, household, and week level along with UPC-specific time (linear) trends.

Two-Part Model

We selected a Two-Part Model (TPM) to evaluate the impact of the NuVal scores on the shoppers purchasing decisions for two main reasons. First, the adoption of NuVal labels can not only result in households substituting unhealthy products for healthy options (i.e., decreasing the purchase quantity of lower-scoring products and increasing the purchase volume of higher-scoring products) but it might also induce households to purchase scored products, especially those products with higher scores. Ignoring the effect of NuVal labels on the likelihood of buying a labeled product may mislead the interpretation of their impact on consumer behavior.

A suitable approach for evaluating the impact of NuVal on the two-step household purchasing behavior (i.e., the buying decision and the quantity decision) is the TPM. It was first developed by Cragg (1971) and has been employed to estimate data with a large number of zero food purchases (Duffey et al., 2010, Haines et al., 1988). Although more sophisticated estimation methods have been developed based on the TPM such as the Sample Selection Model, they generate similar results as the TPM model (Yen, 2005).

The second reason for choosing a TPM is that it also allows us to account for zero purchases, an important feature of our household-level data. Our sample consists of households who only bought 13% of the time frozen dinner products every time they visited the store.

We estimate the first part of the TPM as the participation equation and the second part as the purchase quantity decision at the UPC, store, household, and week level as shown in equation (21). A TPM can analyze an outcome with a mixed distribution; for instance, an outcome that displays a distribution of a discrete point-mass variable and a distribution of a continuous variable (Lachenbruch, 2002).

Moreover, our household scanner data consist of a mass of zero purchases in the first part of the distribution (Table 2.2) followed by right-skewed data (Figure 2.3). This non-normal distribution of the data can be accommodated in the conditional part of the TPM.

Following the notation of Shonkwiler and Yen (1999) and Tooze et al. (2002) for a demand system, we define the TPM for one product category as follows:

$$(21)\qquad y_{it}^{*} = f(X_{it}, \beta) + u_{2i} + \epsilon_{it}$$

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

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

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 are suppressed for notational simplicity. The equation system in (21) indicates 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:

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

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:

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

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

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

Quantity Decision

The second part of the TPM is conditional on observing the shopper making non-zero purchases as is specified as follows:

(24) $$V_{hirt} = a_i + a_i * t + a_t + a_r + a_h + b_1\, P_{itr} + b_2\, Adopt_{itr} + b_3\, Adopt_{itr} * Score_i + M'_{itr}\gamma + \ \epsilon_{hitr}$$

where v_{hitr} is the purchase volume (i.e., positive purchases) 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 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 interaction term $a_i * t$ allows labeled and non-labeled products at week *t* to follow different linear trends. All UPCs in our sample eventually received a NuVal score during the study period. Labeled products are products identified by their UPC and that had been assigned the NuVal scores at week *t* by the store that adopted NuVal.

Annual sales of frozen entrees declined by more than 2% from 2010-2014 as consumers' preferences shifted towards natural foods (Strom, 2015); therefore, including time fixed effects in the regression model control for this decline in sales and other time-related factors that affect purchases. In addition, including UPC fixed effects in the model address endogeneity due to the possible correlation between unobserved product characteristics and price when UPC characteristics are not controlled.

While the term $Adopt_{itr}$ captures the average NuVal effect of the NuVal label on the labeled UPCs at the NuVal store, the interaction term $Adopt_{itr} * Score_i$ isolates the effect of nutritional information provided by the experts via the NuVal score from the overall label effect.

One might be concerned that promotional activities by the non-NuVal and NuVal stores could have changed during the study period. For instance, the NuVal store could have increased marketing activities for products with lower scores during the NuVal period; therefore, we include the vector M_{itr} that controls 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.

Participation Decision

The first part of the TPM is specified as in (24) with d_{hirt} , the binary choice variable, as the dependent variable:

(25) $d_{hirt} = a_i + a_i * t + a_t + a_r + a_h + b_1\, P_{itr} + b_2\, Adopt_{itr} + b_3\, Adopt_{itr} * Score_i + M'_{itr}\gamma + \epsilon_{hitr}$

Quantile Estimation

Suppose that Y is a random dependent variable (purchases or some monotone transformation of purchases) with a cumulative distribution function $F(y|x)$=P($Y \leq$ y|x) and x is a vector of covariates. In the quantity decision (second part) of the TPM, the conditional mean function of purchases can be represented as $E(Y|x, d = 1) = x'\beta$, where $d = 1$ indicates non-zero purchases. In contrast, a linear-in-parameters quantile regression can estimate the τ conditional quantile of Y as

(26) $Q_Y(\tau|x) = F^{-1}(\tau) = \inf\{\text{y}: F(y|x) \geq \tau\},\ 0 < \tau < 1$

where τ is the quantile level; $Q_Y(\tau|x)$ is the τth percentile. The quantile regression coefficient $\beta(\tau)$ can be estimated by minimizing the function over β:

(27) $r(\beta) = \sum_{i=1}^{n} \rho_\tau(Y - x'\beta(\tau))$

where the loss function $\rho_\tau(z)$ is defined as $z\big(\tau - I(z < 0)\big)$ and the residuals are defined as $z = Y - x'\beta(\tau) = z$. This loss function assigns a weight of τ to positive residuals and a weight $1 - \tau$ to negative residuals (z).

The main advantage of quantile estimation is that it can estimate the entire conditional distribution of purchases, allowing for the covariate effect β to change with the dependent variable. This approach is considered when describing noncentral positions (e.g., tails) of an outcome distribution is more interesting (Hao and Naiman, 2007). In addition, when the outcome distribution is highly skewed, the conditional mean effect can be challenging to interpret while the median effect can be more informative. For the median or the $50th$ percentile ($i.e., \tau = 0.5\ or\ 0.5\ quantile$), the coefficient $\beta(0.5)$ minimizes the sum of the absolute residuals.

We estimate the quantile regressions employing the smoothing algorithm in SAS. According to Chen (2007), the smoothing algorithm outperforms the simple and interior algorithm for quantile regressions in terms of accuracy and computing speed.

Heterogeneous Consumers' Responses to NuVal

Socio-demographic groups

Educated meal planners are generally more aware of diet-health links than other household members (Variyam et al., 1996), and therefore they are more likely to use nutritional information (Drichoutis et al., 2005, Nayga, 1996). Because time spent at grocery shopping has a positive effect on label use (Drichoutis et al., 2006), it is plausible that time-constraint households have different patterns of information search behavior compared to households that do not face time pressures. For this reason, we expect that employment status, income level, marital status, and children presence could be related to the use of nutritional labeling. Contrary to the impact of education, the effects of working status, income, and household size on food labeling use remain inconclusive (Drichoutis et al., 2006). Empirical evidence indicates that household characteristics including marital status and time spent at work can be important influencing preferences for frozen foods and take-out meals (Harris and Shiptsova, 2007, Hunter and Worsley, 2009). Therefore, we also classified individuals based on these characteristics.

Because NuVal scores summarize the nutrition profile of food products including information on the Nutrition Facts labels, we expect modest or no improvements of the healthfulness of food choices among nutrition label users (e.g., college-educated households) and large effect on no college educated shoppers. Previous observational research showing that less-educated individuals benefit the most from simplified labeling formats supports this hypothesis (Zhu et al., 2015).

We employ separate regression analyses to estimate the impacts of posting NuVal shelf labels (i.e., the NuVal effects) for different demographic groups (e.g., low- and high-income families, families with and without children, families with and without college-educated heads). While one can test heterogeneous effects by employing interaction terms between indicator variables for each household group and the NuVal variables (i.e., $Adopt_{itr}$ and $Adopt_{itr} * Score_i$), we conduct separate regression analyses for three main reasons. First, the large number of parameters in our empirical model based on large-scale purchase data compromises the speed of the estimation process, especially when using non-linear estimation procedures (e.g., logit models).

Second, separate regression analyses allow us to estimate separate estimates for NuVal effects, own-price elasticities, marketing variables without the adding more variables (i.e., interaction terms). Finally, while the advantage of testing heterogeneous effects by employing one regression with interaction terms is that we can test the effect of one variable while controlling for changes in other predictor variables, the ceteris paribus assumption might not hold in some cases (Kennedy, 2005). For instance, in our sample, one would expect that income level will change when a household head retires.

We estimate quantile regressions for each household group, to test these heterogeneous effects across different levels of purchase intensity. To be able to control the effect of other household characteristics (observable and unobservable), we include household fixed effects in all regression models (i.e., TPM and quantile regressions.

Data

We used scanner data of frozen dinner products from the IRI Academic DataSet to track household-level purchases and store-level prices in a small Midwestern city before and after the adoption of NuVal.

Household- and store-level scanner data from six grocery stores in a Midwestern city were used. Only one store adopted NuVal (the NuVal store), and no other stores in the city adopted either Guiding Stars or NuVal labels during our sample period.

Household food purchases by panelists at these retailers were automatically captured at the store checkout. This data collection method reduces the incidence of misreported prices and quantities compared with data collected through in-home scanning (e.g., Nielsen Homescan). We use card panelists 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, were provided for this research. UPC-level NuVal scores for frozen dinner were obtained from NuVal LLC, NuVal's licensing company.

Because our NuVal store adopted NuVal in August 2010, we define September 2010 to December 2011 as the adoption period and January 2010 to August 2010 as the non-NuVal period for our analysis.

Our sample consists of 251 UPCs that were sold during both the NuVal and non-NuVal periods. We used UPCs that exist before and after post label period so we can control for the possibility that post label period the store might have a higher introduction of healthier products (i.e., private brands). The NuVal scores for the products in our regression sample range from 3 to 50, with an average of 20.49.

Differences between the NuVal and Non-NuVal Stores

Table 2.1 provides the average weekly quantity purchased at the NuVal and non-NuVal stores. The first two columns of the table report the mean purchases for all labeled UPCs, regardless of the NuVal scores. In the remaining columns of Table 2.1, we investigate whether the NuVal effect is different for UPCs with higher and lower NuVal scores.

First, to examine the differences between the non-NuVal stores and the NuVal store, column 1 provides the summary statistics for the frozen dinner products purchased in these stores during the non-NuVal period. The summary statistics show that households on average purchased almost similar volume of a UPC per week in the NuVal store and non-NuVal stores during the non-NuVal period (1.07 and 1.04 units). During the study period, while the NuVal store carried 81% of the labeled UPCs of our sample, the non-NuVal stores carried 99%.

Columns 4 to 7 correspond to unit price, advertising, price reduction, and score information of the frozen dinner products at the NuVal and the non-NuVal stores during the NuVal and the non-NuVal periods. The Unit Price column is the price per unit paid by the household at the checkout at the NuVal and the non-NuVal stores during the NuVal and the non-NuVal periods. The subsequent columns indicate whether the product had a coupon or any other advertising (*Ad*), whether the product had a price reduction tag (*PR*), and the NuVal scores (*Score*). Although the NuVal and non-NuVal stores did not carry the same number of UPCs, the prices and NuVal scores were similar during the study period. The prices per equivalised unit charged by the NuVal and the non-NuVal stores during the non-NuVal period were $4.20 and $4.11, respectively. The mean score of the UPCs in the NuVal store and the non-NuVal stores was about 29 (i.e., 20.61 and 20.36, respectively). In terms of retail marketing strategies, the non-NuVal stores advertised UPCs more than the NuVal store (7% vs. 3%); however, they posted less discount price tags than the NuVal store (23% vs. 26%).

Differences in Means

In this section, we compare changes in weekly purchases at a household-level before and after the adoption of NuVal without controlling for covariates. On average, 56.20% of UPCs in our sample are scored more than 21 and 43.80% of UPCs are scored less than 21. We select this cutoff because the mean score is 20.49, therefore 21 is a natural cutoff point for our sample that contains scores ranging from 3 to 50. These percentage values indicate that the distribution of NuVal scores is roughly normal (Figure 2.3).

A comparison of changes in purchases at the NuVal store indicates that there was an increase in purchase volume by 0.18 units of UPCs with scores equal or higher than 21 (Column 3). As expected, there was a decrease in purchase volume by 0.10 units of UPCs with scores lower than 21 (Column 2). The overall volume effect was an increase in purchases by 0.06 units at the NuVal store after the NuVal label adoption (Column 1).

During the same sample period, at the non-NuVal stores, the average volume purchased for each labeled UPC regardless of the score increased by 0.05 units. Thus, the estimated effect, measured as the difference between change in means (at the non-NuVal stores and the NuVal store), suggests that posting NuVal labels increased consumer demand for the labeled UPCs by 0.01 units, or about 0.8% of 1.07 units which is the average volume purchased in the NuVal store before the NuVal labels were adopted. This volume change across all UPCs can be calculated as the average volume change weighted by the percentage values of UPCs higher and lower than 21 (i.e., -0.10 x 43.80% + 0.18 x 56.20%).

There was an increase in the unit prices of frozen dinner products at the non-NuVal and NuVal stores after the NuVal adoption (i.e., by \$0.07 and \$0.11, respectively). Marketing activities also increased at both the NuVal and non-NuVal stores after the NuVal labels were implemented. Overall, advertising increased by about 2 percentage points in all stores and price discount tags increased by 1 percentage point at the non-NuVal stores and by 2 percentage points at the NuVal store.

While the approach to compare of means is informative, it does not consider the effects of observed and unobserved factors related to product, store, and household on frozen dinner consumption. Therefore, in the next section, we use regression analysis to control for these factors and estimate the NuVal label effects across household groups.

Sample Characteristics

Table 2.2 reports the summary statistics for all variables used in the analysis of households purchasing behavior. The summary statistics indicate that only 0.09% of the shopping trips correspond to frozen dinner purchases of UPC *i* in week *t* at store *r*. On average, households bought about 1.07 equivalised units of each UPC per week at a price of 4.22 dollars per unit. Because different UPCs have different package size, units were normalized, specifically, a UPC with a package size of 454 grams is equivalent to 1 equivalised unit. On average, 7% of the UPCs were advertised, and 24% were labeled with a price reduction mark that indicated a price discount of at least 5%.

The summary statistics related to the shoppers' socio-economic status in our sample show that 28% are low-income households and 58% have not attended college. Information corresponding to household composition indicates that 20% have at least one child, 66% are single, and 12% are married. The summary statistics of employment status point out that the head of the household of about 10% of the families was not employed; 24% of these household heads worked full-time, and approximately 8% has a household head that is retired. We classify shoppers as low-income or high-income households based on the 185% of federal poverty guidelines.

Empirical Results

Two-Part Model

The estimation results of the extensive purchase decision (i.e., logistic regressions) and the intensive quantity decision (i.e., conditional regressions) for the whole sample and each household group are presented in Table 2.3.

Extensive Purchase

The regression results for the extensive purchase decision for the whole sample and each household group indicate that the parameter estimates for price (*P*), advertising (*Ad*), and price reduction tag *(PR)* are statistically significant and have the anticipated signs across all models. 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 likelihood of a household making a purchase of this UPC will increase.

The odds ratio estimates are presented in Table in Appendix H. The largest effect of advertising on the likelihood of buying a UPC is among families with married household heads. For these households, the odds of buying a frozen dinner product that is being advertised are 79% higher than the odds of buying a non-advertised product.

Not surprisingly, the effect of a price discount tag on the likelihood of buying a product is largest among low-income households. For these households, the odds of buying a product increase by 96% when the product has a price reduction mark.

The parameter estimate for *Adopt* in the regression for the whole sample is statistically significant. This provides evidence that the NuVal labels influence the household's probability of making a purchase after frozen dinner products were labeled at the NuVal store. However, we do not find evidence of the impact of the NuVal scores (i.e., the interaction term *Adopt*Score*) on the participation decision for the whole sample. On the contrary, separate regression analyses across different demographic groups show that the NuVal scores influenced the participation decision of specific demographic groups.

The results of the regression analyses indicate that NuVal scores increase the likelihood of choosing healthy products over unhealthy ones for some households in our sample. More precisely, NuVal scores increase the probability of buying healthier frozen dinner products for high-income households and families with children. Surprisingly, after controlling for sociodemographic characteristics of households, low-income families, households with married couples, and families whose household heads were unemployed were less likely to choose frozen dinner products with higher scores over lower-scoring products.

For high-income households, the probability increases by 1.2 % for every point increase in NuVal score. In contrast, the probability of making a purchase for low-income households and families with married household heads decrease by 2% and 2.4%, respectively when the NuVal score of a UPC increases by 1 point. For households with children, the probability of making a purchase of a UPC increases by 2.1 % when the NuVal score increases by 1 point. For families with household heads that were unemployed, the probability of buying decreases by 3.4 % for every point increase in the NuVal score.

Using the estimates of the logit model for the whole sample, we can calculate the dollar value of the NuVal label for the average frozen dinner product. The dollar value at the mean score is \$0.18 [i.e., (0.004*21-0.192+0.061)/-0.602]. Because the average unit price is \$4.22, this value represents almost 2% of the price per unit.

Intensive Purchase

The regression results for the intensive purchase decision in Table 2.3 indicate that the scale parameter estimates range between 0.41 and 0.55. This parameter was estimated as the square-root of the normalized Pearson's chi-square. Because the scale estimates across all models are different from 1, selecting a normal distribution (i.e., restricting the scale parameter to be 1) for modeling purchases would be incorrect.

The parameter estimate for price (*P*) is statistically significant and negative across all models, except for one household group. For families with a retired household head, the price estimate is not statistically insignificant. The lack of statically significance of this parameter could be because this group is the smallest in our sample (i.e., 8%). The parameter estimates for the marketing variables (i.e., *PR* and *Ad*) are statistically significant for most the household groups. The parameter estimate for the advertising variable (*Ad*) is statistically significant and positive for all regression models except for three. This parameter is not statistically significant in the regression models for households with no college education, families with a household head that is not employed, and households who have a retired head. The parameter estimates for the price reduction tag variables (*PR)* are positive and statistically significant for all regressions except for the regression models that explain purchases of low-income shoppers, households with a college education, families with children, and retired household's heads. Overall, our results showing that households with a head who is retired are not sensitive to price reductions or promotional activities when shopping frozen dinner products contrast the idea that these households might spend more time searching for deals because of their lower cost of time.

The Table in Appendix H indicates the partial effect estimates for the conditional models. The estimates indicate that high-income households had the larger response to advertising. These households increased purchase volume by 0.22 units when the product was advertised. On the contrary, the lowest effect of advertising on purchase volume was found among families who have a retired household head. These shoppers increased purchases by 0.07 units when the product was advertised. Interestingly, families with both household heads working full time are more responsive to price discounts when deciding how many units of frozen dinner they will buy. This demographic group increased volume by 0.16 units when the product has been assigned a price discount tag. Shoppers who attended to college increased purchase volume by 0.05 units. This estimate corresponds to the lowest effect of a price cut tag on purchase volume.

We found that NuVal scores increase purchase volume for high-scoring products and decrease it for low-scoring products among low-income families, single shoppers, households with children, and shoppers without any college education (i.e., the parameter estimate for *Adopt*Score* is positive and statistically significant for these shoppers).

Previous work indicated that although simplified labeling formats help consumers to identify healthy items, they have little influence in their food choices (Borgmeier and Westenhoefer, 2009). In contrast, this study indicates that simplified labels based on nutritional scoring systems influence frozen dinner entrees purchases of low-income households, families with a household head that did not attend college, families with children, and families with household heads that are single.

Purchasing Patterns

Table 2.4 indicates features of the demand for frozen dinner across the household groups. The first two columns indicate the criteria used to classify households (Category) and household groups for each category, respectively. Column 3 indicates the number of UPCs purchased for each household group. The largest heterogeneity (at UPC level) in product selection was found among high-income households and families with no children. These household groups purchased almost all the UPCs in our sample (i.e., 241 out of 242). The smallest heterogeneity in preferences corresponds to households who are retired, who purchased only 81% of the UPCs (i.e., 197 out 242). These households represent only 8% of the sample which might explain the small variability in preferences for frozen dinner at UPC level. It is also possible that these shoppers have well-defined preferences for particular brands of frozen dinner entrees.

Column 4 indicates the frequency that each household group in the sample purchased a specific UPC when visiting the store. The data at UPC level has a high degree of zero purchases (at weekly frequencies) for all households (about 99%). The smallest and largest frequency indicate that for every visit to the store, households with full-time employed heads and low-income families purchased a specific UPC 0.077% and 0.095% of the time, respectively. Similar conclusions are obtained if we aggregate the purchase incidence (at UPC level) across UPCs at week and store level (column 5). For every shopping trip, full-time employed households buy approximately 8% of the time, while low-income families buy frozen dinner products 14% of the time.

Column 6 indicates the average purchase volume of frozen dinner. On average, full-time employed households buy a larger volume of frozen dinner than unemployed and retired households (1.77 vs. 1.73 units). Previous studies indicate that households who face time constraints such as families of employed parents were more likely to eat take-out foods or prepared entrees in their family meals (Devine et al., 2009, Hunter and Worsley, 2009, Park and Capps, 1997). Contrary to this, household heads that work full time in our sample did not buy with higher frequency frozen dinner compared with households who spend less time at work such as retired shoppers and unemployed households (7.6% vs. 9.32% and 8.85% of the time, respectively). While apparently, the fact that households with tight schedules (e.g., full-time employees) have lower preferences for frozen dinner contradicts the theory in the previous section, it is probable that these shoppers prefer take-out meals over frozen meals when looking for convenience foods.

On average, households with children consume more often and buy more units of frozen meals than families without children (13.20 % vs. 12.37 % and 2.01 vs. 1.80 units). Single-person households buy less often frozen dinner products but a larger volume compared to married households (11.74% vs. 13.81% and 1.89 vs. 1.59 units). The purchasing behavior of single and married shoppers in our sample support previous evidence indicating that they spend more in ready foods (canned, dry, frozen, ambient, and chilled meals) compared to those who are married (Harris and Shiptsova, 2007).

Overall the purchasing preferences of the households in our sample indicate that there is some evidence that time-constrained households buy frozen dinner more frequently compared to shoppers who have more leisure time (i.e., married vs. single shoppers and families with children vs. families without children). Similarly, we found evidence that households with less time available usually buy a higher volume of these products than shoppers who do not face time constraints (i.e., families with children vs. families without children and full-time employed vs. unemployed and retired shoppers).

The last two columns indicate the own-price elasticity values for frozen dinner. The (conditional) elasticity values derived from the conditional analysis of the second part of the TPM (i.e., quantity decision) indicate that all shoppers have a price-inelastic demand (i.e., values ranging from -0.56 to -0.13). The households who are more price sensitive are unemployed shoppers with a demand elasticity of -0.56, while families whose households' heads are married have a less price-responsive demand (-0.13).

We also calculated the unconditional price elasticities based on the parameter estimates of the quantity and participation decisions of the TPM. Like the findings of the conditional elasticity analysis, the results of the unconditional elasticities indicate that the highest price sensitivity is among unemployed households (-2.95).

Interestingly, our results indicate that the unconditional demand for frozen dinner by households with low-income families is more inelastic than the demand by high-income shoppers. However, the results of the conditional analysis indicate that the (conditional) demand of low-income households is more elastic than the (conditional) demand of high-income shoppers.

Another key difference between the conditional and unconditional estimation in terms of elasticities is that the former indicates that frozen dinner products are highly elastic, as the elasticities range from -2.95 to -1.93, while the latter indicates that frozen dinner is inelastic with values ranging from -0.56 and -0.13. Zhen et al. (2013) employed a censored demand system that addresses price endogeneity and found an elasticity value of -0.765 for this food category.

Similar to previous research that reports a more inelastic demand for food and non-food products of full-time employed shoppers for food and non-food products compared with the average buyer (Ainslie and Rossi, 1998), we found that the demand (conditional and unconditional) of full-time employed shoppers is more inelastic compared to the demand of unemployed shoppers. Related to this study, our results also indicate that retired household heads have the smallest unconditional demand elasticity (-1.93) for these foods.

Heterogeneous Consumers' Responses to NuVal

NuVal Effects on Purchases across Socio-demographic Groups

Because the NuVal labels not only increase the purchase volume of healthier products but also affect the shoppers' likelihood of buying products with higher scores, estimating the impact of the labels based exclusively on the analysis of conditional purchase data will only capture the partial impact of the NuVal labels on consumer purchasing behavior. To consider changes in these probabilities in the analysis of purchasing behavior, we estimate the unconditional NuVal effects for each household group by using the TPM parameter estimates.

For comparison purposes, we also calculate the conditional NuVal effects using only the parameter estimates of the quantity decision models including the scale parameter estimate. We estimate the unconditional and conditional NuVal effects for UPCs that do not have a price mark-down or a discount coupon (i.e., *PR* and *Ad* are 0) at three different values of score (i.e., max, mean, and min) and at the mean unit price (i.e., 21 and $4.22). Calculations of the conditional and unconditional NuVal effects for a log-normal distributed variable are shown in the Appendix D section.

Conditional NuVal Effects

Table 2.5 indicates that the conditional NuVal effects expressed in units and percentage change with respect to the predicted purchase volume before the NuVal adoption. The NuVal effects are reported for UPCs with three different values of NuVal scores (min 3, mean 21, and max 50). We only report the NuVal effects for those household groups to whom the NuVal scores influence the purchase volume (i.e., parameter estimates of the NuVal variables are statistically significant for the quantity decision in Table 2.3). Table 2.5 indicates that families with children, low-income shoppers, households with no college education, and single shoppers increase their purchases for higher-scoring products by 156.4%, 55.2, 54.5%, and 46.8%, respectively.

The largest NuVal effect for a UPC with the highest NuVal score (i.e., 50) among families with children indicate that these shoppers experienced the largest improvement in their food choices after the NuVal labels were adopted. The largest decrease in purchases of a UPC with a score of 3 occurred among low-income shoppers. These households decreased on average their purchase volume of this UPC by almost 22%. Our results that indicate that low-income households experience the largest response to the labels support findings of a previous empirical study that indicates that female shoppers with high income do not respond to FOP labels Zhu et al. (2015).

Unconditional NuVal Effects

We estimated the unconditional NuVal effects to account for changes in shoppers' likelihood of buying UPCs with higher and lower NuVal scores. Because NuVal scores influence both the purchase volume and the probability of buying frozen dinner products (i.e., extensive and intensive purchase decisions) for only low-income households and families with children, we only report the NuVal effects for these shoppers. Comparing the NuVal effects from the conditional and unconditional analyses, we can observe that the unconditional NuVal effects for low-income and families with children in Table 2.6 are different in magnitude compared with the conditional NuVal effects in Table 2.5.

Like the conditional analysis, the largest increase in purchases of a UPC scored 50 occurred among shoppers with children (148%). However, the conditional analysis estimates a smaller NuVal effect (47%). These findings indicate that the conditional analysis understates the effect of the label for this household group.

For low-income households, contrary to the result of the conditional analysis that indicates that these shoppers increased purchases by 31% for the highest-scoring product (NuVal score of 50), our results of the unconditional analysis indicate that they decreased purchases for this UPC by 28%. This difference in the NuVal effect between conditional and unconditional purchases for low-income households is mainly attributed to the fact that the parameter estimate for the NuVal variable (i.e., *Adopt*Score*) in the logit model is negative while in the conditional model the parameter estimate is positive (Table 2.3).

Although we found differences in the purchase responses to the NuVal labels between the unconditional and conditional analysis, results from both analysis indicate that the largest improvement in food choices happened among those shoppers who have at least one child.

NuVal Effects on Purchases across Different Levels of Purchasing Intensity

It is possible that shoppers' responses to the NuVal scores change across the distribution of purchasing levels (i.e., the effect of NuVal scores is different between heavy- and light-user shoppers). To test for heterogeneous NuVal effects across different levels of purchasing intensity, we estimate quantile regressions. We report the estimation results only for those households, whose mean NuVal effects were statistically significant in the conditional part of the TPM (i.e., low-income shoppers, families with children, single buyers, and no college educated household heads). Figure 2.4 describes the median and the frequency of (conditional) purchases for these 4 household groups. The distributions indicate that purchases are highly-skewed to the right. The mean and median values of purchases for all households are approximately 1.1 and 0.7, respectively. This indicates that the quantile regression estimation for the quantiles that are lower than the median values might be challenging because the small variability of the conditional outcome which can prevent us to estimate the effect of the covariates.

Figure 2.5 and Figure 2.6 indicate the parameter estimates for the unit price and the NuVal variable (i.e., *Adopt*Score*), respectively. The quantile regressions were estimated between the 10th and 90th percentiles. The parameter estimates for the lowest percentiles (i.e., between the 10th and 40th percentiles for low-income shoppers and families with children, and the 10th and 60th percentiles for shoppers without a college degree and single buyers) are close to zero and are not statistically significant.

As expected, the parameter estimates for the unit price are negative across all conditional quantiles for all four household groups. The downward sloping curve for the price effects across quantiles indicates that, in general, shoppers who buy a larger volume of frozen dinner (i.e., heavy users) are more price sensitive than those who buy smaller volume (i.e., light users). Our finding is consistent with the information-theoretic approach described by Kim and Rossi (1994). It is possible that consumers with preferences for a higher volume might have a better understanding of the distribution of prices, and therefore they are more sensitive to price cuts.

Figure 2.5 also reveals that the heterogeneous effects of price across quantiles for some households are non-linear. For shoppers with no college education, the sensitiveness to price changes increases in a linear manner as purchasing intensity increases. For the other three household groups, price responsiveness increases up to the 80th percentile and it is constant (i.e., low-income shoppers) or decreases (i.e., shoppers with children) between the 80th and 90th percentiles.

The quantile regression results for the NuVal scores (i.e., *Adopt*Score*) in Figure 2.6 indicate that the response to NuVal scores of shoppers without college education increases above the 80th percentile and is statistically different from zero at the 90th percentile. For families with at least one child, the NuVal effects increase between the 40th and 60th percentile and remain approximately at the same value (0.01) between the 60th and 90th percentiles with small fluctuations. For single-person households and low-income shoppers, the effects of NuVal scores are increasing above the 80th percentile; however, the NuVal effects across the distribution of purchases are not statistically different from zero. The lack of statistical significance of NuVal scores for these household groups contradicts the findings of the conditional analysis (second part) of the TPM. We found that the parameter estimate for the NuVal variable (i.e., *Adopt*Score*) for these shoppers was positive and statistically significant in the conditional part of the TPM. Differences in the distribution assumption of each methodology might explain these contrasting results. While the log-normal distribution was chosen to fit purchases in the conditional part of the TPM., quantile regressions are estimated with no distribution assumptions (Koenker, 2005). Overall, the impact of NuVal scores is larger in the upper quantiles for households with children (i.e., above 0.8 quantiles) and no college household heads (i.e., above 0.6 quantiles).

Finally, we compare the heterogeneous NuVal effects from the conditional quantile regressions with the NuVal effects from the conditional part of the TPM for households with children families with household heads without a college education (the only two household groups which NuVal effect was statistically significant in the quantile estimation). Table 2.7 indicates that for households with children the parameter estimates for the NuVal variable (*Adopt*Score*) are between 0.009 (the 80th percentile) and 0.013 (70th percentile). The parameter estimate in the conditional analysis for this variable is 0.013. For no college educated shoppers, the parameter estimate is only statistically significant for the 90th percentile (i.e., 0.009). This value is close to the estimate of the conditional analysis (i.e., 0.008).

Table 2.7 also indicates the (conditional) NuVal effects at the means for a UPC with the maximum NuVal score and no marketing efforts [i.e., *TE* (max score)]. For households with children, the highest increase in purchases because of NuVal adoption was found at the 70th percentile (41%). In contrast, the (conditional) NuVal effect in the second part of the TPM was 47%. For non-college educated shoppers, only purchases at the 90th percentile were affected by the NuVal labels adoption. The increase in purchases at this percentile was 35%, which is almost similar to the increase in purchases computed in the conditional TPM analysis (36%).

As a whole, the quantile regression results indicate that heavy users (i.e., high-volume shoppers) have a greater response to the NuVal scores. It is probable that light users might not be concerned about the nutritional value of products that they purchase in small volume, while shoppers who purchase a larger volume of these are more conscious that their choices will have a greater impact on the nutritional quality. Alternatively, low-volume shoppers are less concerned about their food choices because the magnitude of their purchases will not have a significant impact on their diet. Our results do not show support for previous work that indicates that shoppers with occasional purchases benefit the most from simplified labeling systems (Zhu et al., 2015).

Our findings have broader implications for applications of consumer demand models to food and nutrition policy research. Our results suggest that the practice of estimating the average demand should be complemented with a deeper analysis of the demand across the demand distribution by relaxing certain distributional assumptions of the outcome.

Conclusions and Implications

Because the proliferation of FOP labels on processed foods that might lead to consumers' confusion, the possibility to develop a simplified, standardized, and science-based nutrition symbol has been discussed by policymakers in the U.S. and worldwide (FDA and WHO). Furthermore, because the definition of nutrition claims remains vague, the FDA is considering redefining "healthy". Given these potential changes in labeling policy, examining the effectiveness of current summary nutrition labeling systems at improving food choices is critical. Purchase data when a supermarket's voluntary adoption of NuVal— a 1 to 100 numeric summary shelf label system was employed to estimate a Two-Part Model (TPM) and quantile regression to identify the effect of the NuVal labels on consumer purchasing decisions for frozen dinner products. Our main findings are as follow:

First, the TPM estimation results for different household groups indicate that the adoption of the NuVal labels increases the purchase volume of healthier frozen dinner and influences the likelihood of buying healthier frozen dinner for households with low-income and shoppers who have with at least one child. Employment status did not influence the effect of NuVal indicating that there is no evidence that NuVal scores benefit more to time-constrained households.

Second, the results of the quantile regressions for purchase volume indicate that posting the NuVal labels has a greater impact on the food choices of heavy users compared with light users among shoppers without a college education and households with children. This points out that those households with a higher preference for frozen dinner have greater improvements on the healthfulness of their choices compares to consumers with lower preferences.

Third, the lack of parallel of the results between the quantile regressions and the conditional part of the TPM indicate that assessing shoppers' choices based on predicting the average demand assuming certain distributions might not be enough to understand the impact of the NuVal scores. A closer examination of the demand across the distribution of purchases can provide a better understanding of shoppers' responses to the NuVal scores, especially when purchase data is highly skewed.

Our findings suggest that providing interpretative summary nutrition information about the overall nutritional value of food products can be an effective way to improve consumer choices for only some households.

Table 2. 1 Difference between Non-NuVal and NuVal Stores

		1 Volume (242 UPCs)			2 Volume & Score<21 (106 UPCs)			3 Volume & Score>=21 (136 UPCs)			4 Unit Price		5 Ad		6 PR		7 Score
*Period		Non-NuVal	NuVal	Change	Non-NuVal	NuVal	Change	Non-NuVal	NuVal	Change	Non-NuVal	NuVal	Non-NuVal	NuVal	Non-NuVal	NuVal	
Non-NuVal Stores	Mean	1.04	1.09	0.05	1.03	1.08	0.05	1.06	1.11	0.05	4.20	4.27	0.07	0.09	0.23	0.24	20.61
(239 UPCs)	S.D.	0.96	1.00		0.84	0.90		1.03	1.06		1.49	1.50	0.25	0.28	0.42	0.43	8.31
NuVal Store	Mean	1.07	1.13	0.06	1.15	1.05	-0.10	1.01	1.19	0.18	4.11	4.22	0.03	0.04	0.26	0.28	20.36
(197 UPCs)	S.D.	1.22	1.09		1.11	0.73		0.82	0.92		1.48	1.49	0.16	0.21	0.44	0.45	8.59
Difference		0.03	0.03		0.12	-0.03		-0.04	0.08								
Effect (Volume)				0.01			-0.15			0.13							
Effect (%)				0.80%			-13.32%			12.50%							

Note: S.D. indicates standard deviation. * Because UPCs were scored by the NuVal LLC on different dates, UPCs scored after the NuVal store adopted the labels for the first time have unique NuVal and non-NuVal periods.

Table 2. 2 Summary Statistics for the Regression Analysis

Category	Variable	Description	Mean	S.D.
		Dependent Variables		
Participation Decision	d	d=1 if UPC *i* was purchased in time *t* from retailer *r* by household *h*, 0 otherwise	0.0009	0.030
Quantity Decision	v	Purchase volume measured as the number of equivalised units purchased of UPC *i* in time *t* from retailer *r* by household *h*	1.074	1.023
		Explanatory Variables		
Product Characteristics				
NuVal Label	Adopt	Adopt=1 if UPC *i* had been assigned a score during the NuVal period at the NuVal store at time t, 0 otherwise	0.099	0.299
	Score	NuVal Score of UPC *i*	20.492	8.342
Marketing	P	Price per equivalized unit	4.218	1.494
	Ad	Ad =1 if coupon or if any advertising sign, 0 otherwise	0.070	0.255
	PR	Price reduction flag= 1 if Total Price Reduction is 5% or greater, 0 otherwise	0.243	0.429
Household Characteristics				
Socioeconomic Status	Low Inc	Low Inc=1 if low-income household according to the FPG, 0 otherwise	0.281	0.450
	No College	No College=1 if household heads have not attended college, 0 otherwise	0.575	0.495
Household Composition	Single	Single=1 if household head is single, 0 otherwise	0.660	0.474
	Married	Married=1 if household head is married with no children, 0 otherwise	0.115	0.319
	Children	Children=1 if household has children, 0 otherwise	0.201	0.401
Employment Status	Full Time	Full Time=1 if household heads have full-time job (>35 hours per week)	0.221	0.415
	Not Employed	Not Employed=1 if household head is not working	0.102	0.303
	Retired	Retired=1 if household head is retired	0.084	0.278
Stores				6
Weeks (2010-2011)				103
UPC				242
Households				1495

Note: S.D. indicates standard deviation

Table 2. 3 Estimation Results

	Whole Sample						Low Inc						High Inc					
	Participation Decision			Quantity Decision			Participation Decision			Quantity Decision			Participation Decision			Quantity Decision		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
P	-0.602	***	0.014	-0.071	***	0.02	-0.523	***	0.027	-0.094	***	0.032	-0.632	***	0.017	-0.060	***	0.022
Adopt	-0.192	**	0.084	-0.0404		0.07	0.348	**	0.156	-0.457	***	0.145	-0.383	***	0.100	-0.002		0.079
Adopt*Score	0.004		0.004	0.0043		0	-0.020	***	0.007	0.015	**	0.006	0.012	***	0.004	0.004		0.004
Ad	0.426	***	0.023	0.0972	***	0.03	0.337	***	0.043	0.094	***	0.036	0.455	***	0.028	0.113	***	0.035
PR	0.613	***	0.021	0.0434	*	0.02	0.674	***	0.038	0.030		0.037	0.592	***	0.025	0.060	**	0.025
NuVal Store	0.061	**	0.026	0.6798	***	0.2	-0.092	*	0.050	0.720	**	0.312	0.112	***	0.030	0.768	***	0.186
Scale				0.5579						0.524						0.553		
Households	1495			1495			420			420			1075			1075		
UPC	242			242			233			233			241			241		
QIC/AIC	266140			19865.4			77654			6163			187690			14246		
N	21877985			19190			6005262			5709			15884163			13481		
	No College						College						Married					
	Participation Decision			Quantity Decision			Participation Decision			Quantity Decision			Participation Decision			Quantity Decision		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
P	-0.597	***	0.019	-0.090	***	0.021	-0.615	***	0.022	-0.078	***	0.023	-0.462	***	0.042	-0.031		0.032
Adopt	-0.045		0.118	-0.090		0.116	-0.267	**	0.120	-0.062		0.080	0.335		0.310	-0.025		0.174
Adopt*Score	0.0003		0.005	0.008	*	0.005	0.005		0.005	0.006		0.004	-0.024	*	0.014	0.003		0.008
Ad	0.459	***	0.030	0.046		0.028	0.382	***	0.037	0.137	***	0.041	0.583	***	0.069	0.171	***	0.059
PR	0.605	***	0.027	0.053	**	0.024	0.616	***	0.032	0.039		0.031	0.600	***	0.061	0.126	***	0.037
NuVal Store	0.013		0.036	0.468	*	0.245	0.101	***	0.035	0.350		0.266	0.088		0.079	-0.171		0.277
Scale				0.542						0.524						0.407		
Households	859			859			859			859			172			172		
UPC	236			236			236			236			203			203		
QIC/AIC	156383			11923			110248			8511			31473			2434		
N	12637251			11291			9240734			7899			2391432			2342		

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

Table 2. 3 Estimation Results (Continued)

	Single						Children						No Children					
	Participation Decision			Quantity Decision			Participation Decision			Quantity Decision			Participation Decision			Quantity Decision		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
P	-0.600	***	0.018	-0.052	**	0.022	-0.581	***	0.032	-0.067	**	0.031	-0.606	***	0.016	-0.075	***	0.021
Adopt	-0.283	***	0.099	-0.064		0.070	-0.435	***	0.158	-0.258	*	0.133	-0.106		0.099	0.146	**	0.068
Adopt*Score	0.007		0.004	0.006	*	0.003	0.021	***	0.007	0.013	**	0.005	-0.002		0.004	-0.003		0.003
Ad	0.435	***	0.029	0.107	***	0.036	0.406	***	0.052	0.096	*	0.050	0.429	***	0.026	0.105	***	0.033
PR	0.661	***	0.026	0.055	**	0.027	0.584	***	0.044	-0.002		0.040	0.627	***	0.023	0.072	***	0.027
NuVal Store	0.091	***	0.031	0.525	***	0.190	0.192	***	0.054	0.840	***	0.184	0.015		0.029	0.683	***	0.228
Scale				0.522						0.528						0.547		
Households	986			986			301			301			1194			1194		
UPC	240			240			230			230			241			241		
QIC/AIC	169929			12787			55451			4252			209743			15929		
N	14870544			12139			4349674			3993			17528311			15197		
	Full Time						Retired						Not employed					
	Participation Decision			Quantity Decision			Participation Decision			Quantity Decision			Participation Decision			Quantity Decision		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
P	-0.466	***	0.034	-0.060	*	0.033	-0.42	***	0.05	-0.05		0.03	-0.578	***	0.046	-0.132	***	0.050
Adopt	-0.411	*	0.217	-0.145		0.149	0.09		0.42	-0.21		0.33	0.615	**	0.271	0.115		0.164
Adopt*Score	0.013		0.010	0.010		0.007	-0.01		0.02	0.01		0.01	-0.034	***	0.012	-0.009		0.008
Ad	0.570	***	0.053	0.086	*	0.044	0.45	***	0.09	0.10		0.06	0.379	***	0.075	-0.017		0.051
PR	0.641	***	0.046	0.102	***	0.035	0.58	***	0.07	0.08		0.05	0.599	***	0.066	0.127	***	0.049
NuVal Store	0.034		0.048	0.516		0.351	0.02		0.08	0.38		0.27	0.130	*	0.069	0.500		0.358
Scale				0.549						0.41						0.450		
Households	330			330			126			126			126			126		
UPC	214			214			197			197			197			197		
QIC/AIC	53632			4169			22142			1552			27375			1964		
N	5106026			3907			1850494			1621			2273485			2031		

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

Table 2. 4 Purchasing Patterns across Households

Category	Household Group	UPC	% Buy		Volume	Elasticity	
			UPC level	Across UPCs	Across UPCs	Conditional	Unconditional
Whole Sample	Whole Sample	242	0.088%	12.535%	1.843	-0.300	-2.820
Socioeconomic Status	Low Income	233	0.095%	13.522%	1.817	-0.396	-2.564
	High Income	241	0.085%	12.163%	1.853	-0.253	-2.905
	No College	236	0.089%	12.666%	1.777	-0.380	-2.871
	College	239	0.085%	12.356%	1.940	-0.328	-2.904
Household Composition	Children	230	0.092%	13.200%	2.013	-0.283	-2.710
	No Children	241	0.087%	12.374%	1.801	-0.315	-2.853
	Married	203	0.098%	13.807%	1.588	-0.129	-2.065
	Single	240	0.082%	11.743%	1.886	-0.221	-2.710
Employment	Full-Time Employed	214	0.077%	10.740%	1.767	-0.254	-2.190
	Not Employed	215	0.089%	12.761%	1.725	-0.556	-2.950
	Retired	197	0.088%	12.759%	1.725	-0.194	-1.930

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

Table 2. 5 Conditional NuVal Effects on Purchases across Demographic Groups

Category		Group	NuVal Effects (Units)			NuVal Effects (% Change)		
			Min Score	Mean Score	Max Score	Min Score	Mean Score	Max Score
Socioeconomic Status	**	Low Income	-0.536	-0.221	2.083	-33.81%	-13.92%	31.46%
	*	No College	-0.081	0.100	1.716	-6.37%	7.94%	35.73%
Household Composition	**	Children	-0.395	0.026	2.954	-19.68%	1.32%	47.29%
	*	Single	-0.069	0.100	1.966	-4.47%	6.43%	26.66%

Note: The effects of NuVal labels (i.e., NuVal Effects) at three levels of NuVal scores 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. The statistical significance of the interaction term (*adopt*score*) in the conditional model is expressed as *** p<0.01, ** p<0.05, * p<0.1

Table 2. 6 Unconditional NuVal Effects on Purchases across Demographic Groups

Category		Group	NuVal Effects (Units)			NuVal Effects (% Change)		
			Min Score	Mean Score	Max Score	Min Score	Mean Score	Max Score
Socioeconomic Status	***	Low Income	-0.021	-0.027	-0.041	-14.34%	-18.97%	-28.06%
Household Composition	***	Children	-0.082	-0.004	0.281	-43.00%	-2.02%	147.99%

Note: The effects of NuVal labels (i.e., NuVal Effects) at three levels of NuVal scores 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. The statistical significance of the interaction term (*adopt*score*) in the logistic model is expressed as *** p<0.01, ** p<0.05, * p<0.1

Table 2. 7 Quantile Regression Results across Demographic Groups

		Low income		Children		No College		Single	
Quantile	Variable	Parameter	S.E.	Parameter	S.E.	Parameter	S.E.	Parameter	S.E.
0.6	P	-0.011	0.014	-0.038 **	0.017	0.000	0.002	0.000	0.003
	Adopt	-0.019	0.067	-0.246 ***	0.084	0.000	0.004	0.000	0.003
	Adopt*Score	0.000	0.003	0.010 ***	0.003	0.000	0.000	0.000	0.000
	Ad	0.015	0.022	0.072 **	0.031	0.000	0.003	0.000	0.006
	PR	0.003	0.015	-0.005	0.021	0.000	0.002	0.000	0.003
	NuVal Store	1.4533 ***	0.5545	1.0734 ***	0.336	1.2528 ***	0.3402	0.5596 ***	0.1956
	L.E. (max score)			31%					
0.7	P	-0.048 ***	0.016	-0.044 **	0.019	-0.011	0.008	-0.018 **	0.008
	Adopt	-0.055	0.082	-0.281 ***	0.093	0.003	0.041	-0.006	0.035
	Adopt*Score	-0.001	0.004	0.013 ***	0.004	0.000	0.002	0.000	0.001
	Ad	0.075 ***	0.025	0.071 **	0.033	0.019	0.017	0.038 **	0.015
	PR	0.022	0.021	-0.001	0.023	0.010	0.012	0.021 *	0.011
	NuVal Store	0.8314	0.6058	1.2418 ***	0.4048	1.1455 ***	0.2895	0.6716 **	0.2751
	L.E. (max score)			41%					
0.8	P	-0.053 ***	0.019	-0.048 ***	0.018	-0.030 ***	0.011	-0.030 ***	0.011
	Adopt	0.022	0.096	-0.178 *	0.097	0.017	0.069	-0.005	0.055
	Adopt*Score	-0.004	0.004	0.009 **	0.004	-0.002	0.003	0.001	0.003
	Ad	0.120 ***	0.029	0.059 *	0.033	0.084 ***	0.022	0.055 ***	0.018
	PR	0.057 **	0.025	0.023	0.025	0.080 ***	0.018	0.049 ***	0.015
	NuVal Store	0.527	0.655	1.564 ***	0.488	1.004 ***	0.303	0.737 **	0.341
	L.E. (max score)			28%					
0.9	P	-0.053 **	0.022	-0.030	0.021	-0.054 ***	0.013	-0.055 ***	0.012
	Adopt	-0.061	0.104	-0.248 **	0.106	-0.144 *	0.087	-0.031	0.073
	Adopt*Score	-0.0001	0.005	0.010 **	0.004	0.009 **	0.004	0.003	0.003
	Ad	0.106 ***	0.032	0.046	0.038	0.076 ***	0.024	0.045 **	0.022
	PR	0.074 **	0.029	0.068 **	0.030	0.111 ***	0.021	0.050 ***	0.018
	NuVal Store	0.384	0.654	2.219 ***	0.394	1.081 ***	0.297	0.995 ***	0.347
	L.E. (max score)			30%		35%			

Note: S.E. denotes standard error. *** p<0.01, ** p<0.05, * p<0.1. L.E. (max score) stands for label effect at the maximum NuVal score

Figures

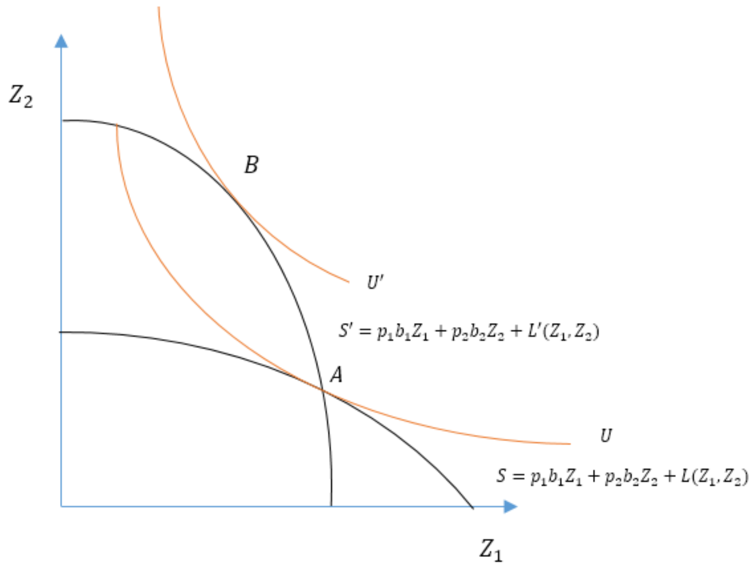

Figure 2. 1 Utility Maximization and Opportunity Cost

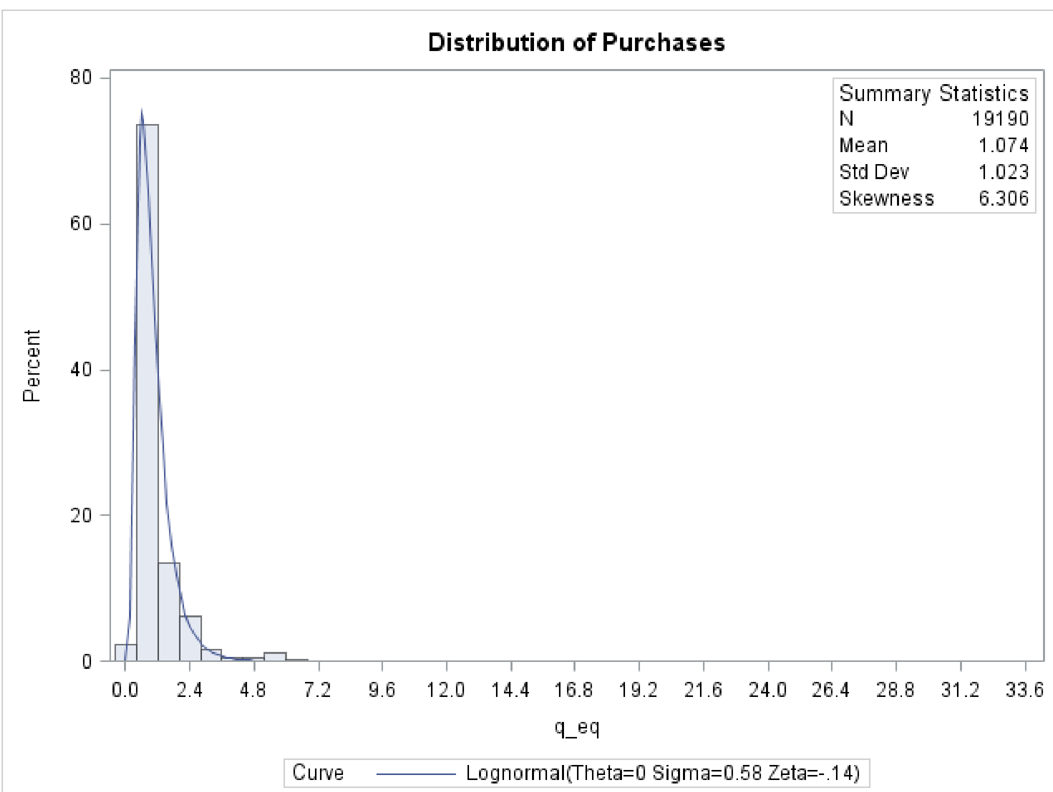

Figure 2. 2 Distribution of Weekly Purchases

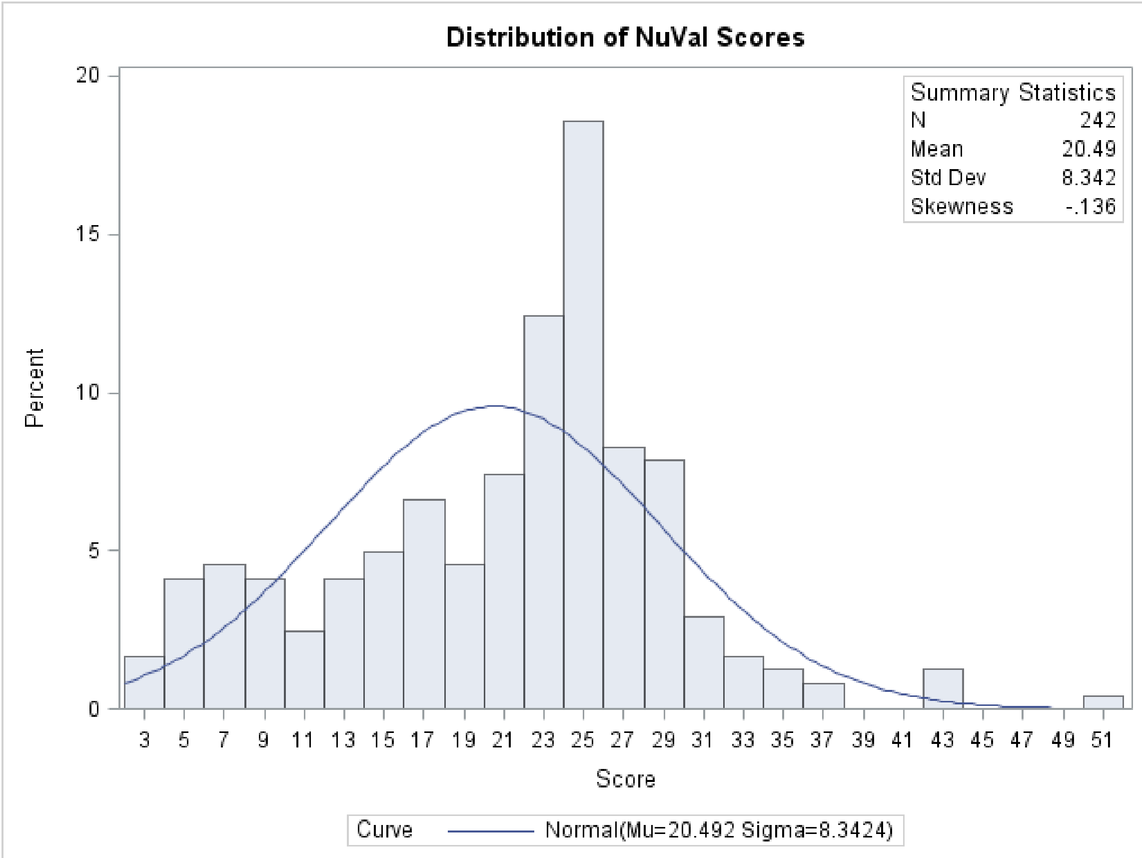

Figure 2. 3 Distribution of NuVal Scores of the UPCs in the Sample of Frozen Dinner Products

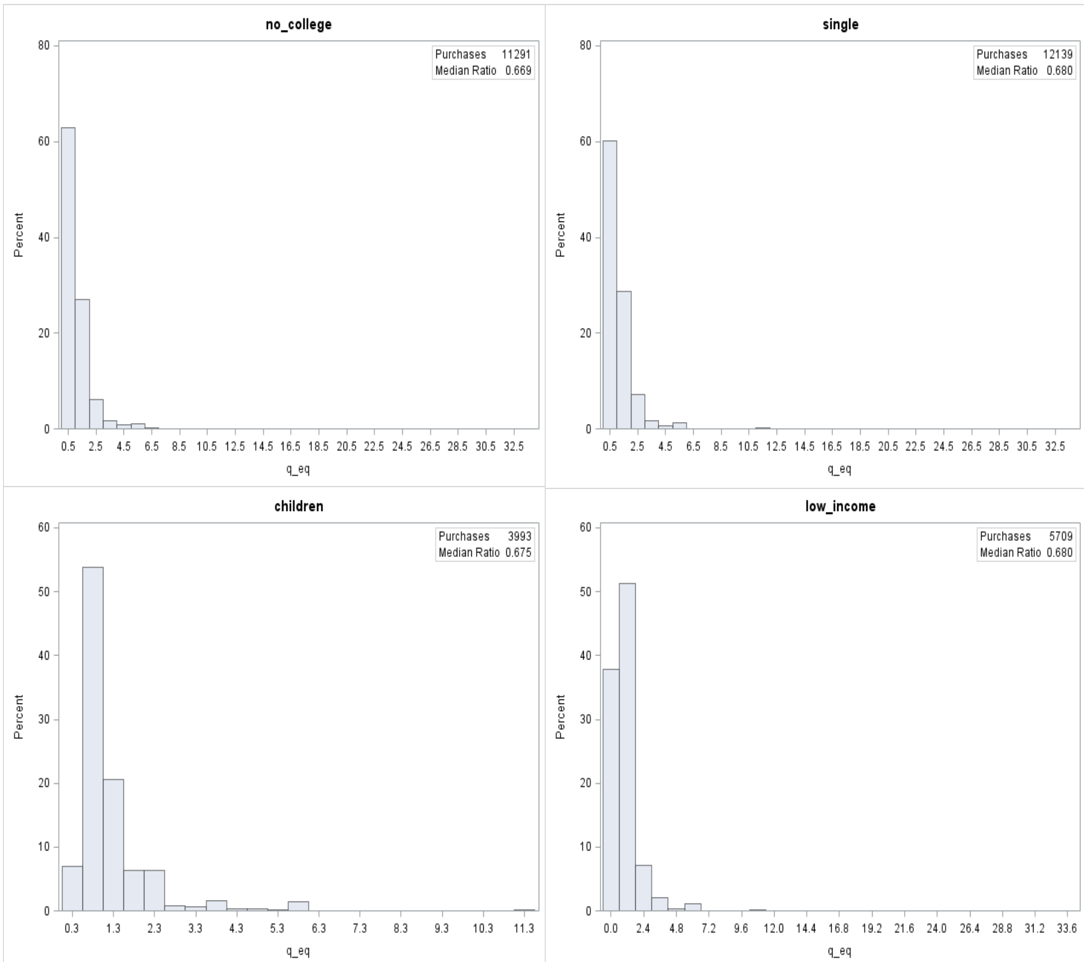

Figure 2. 4 Distribution of Conditional Purchases across Household Groups: no college households (no_college), single-person households (single), families with children (children) and low-income shoppers (low_inc)

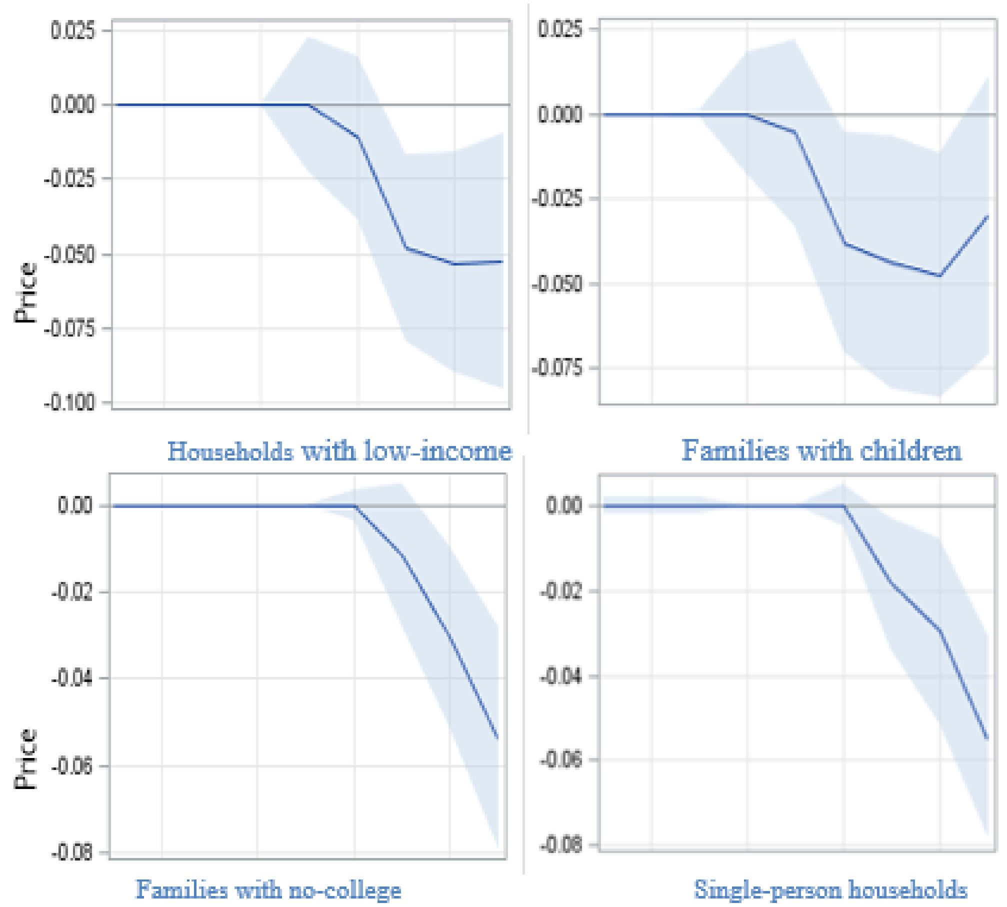

Figure 2. 5 Quantile Regression Results for Price

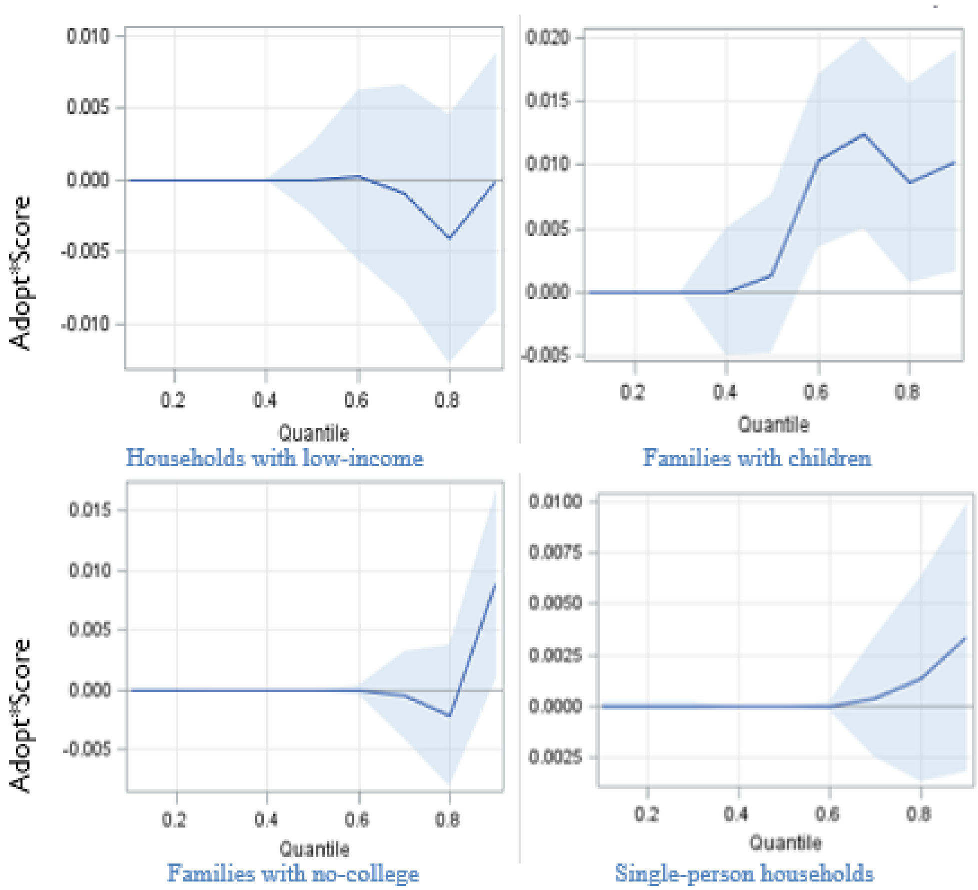

Figure 2. 6 Quantile Regression Results for the Effects of NuVal Scores

REFERENCES

Aguiar, M., and E. Hurst. "Measuring trends in leisure: The allocation of time over five decades." National Bureau of Economic Research.

Ainslie, A., and P.E. Rossi. 1998. "Similarities in choice behavior across product categories." *Marketing Science* 17:91-106.

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

Armstrong, K. 2010. *Stumped at the Supermarket: Making Sense of Nutrition Rating Systems*: Public Health Law & Policy.

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.

Becker, G.S. 1965. "A Theory of the Allocation of Time." *The economic journal*:493-517.

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.

---. 2010. "Do positive nutrition shelf labels affect consumer behavior? Findings from a field experiment with scanner data." *American Journal of Agricultural Economics*:aaq104.

Berry, C., A. Mukherjee, S. Burton, and E. Howlett. 2015. "A COOL Effect: The Direct and Indirect Impact of Country-of-Origin Disclosures on Purchase Intentions for Retail Food Products." *Journal of Retailing*.

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.

Borgmeier, I., and J. Westenhoefer. 2009. "Impact of different food label formats on healthiness evaluation and food choice of consumers: a randomized-controlled study." *BMC public health* 9:184.

Bowers, D. 2000. "Cooking trends echo changing roles of women." *Food Review* 23:23-29.

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., M.J. Sweeney, J. Sobal, D.R. Just, H.M. Kaiser, W.D. Schulze, E. Wethington, and B. Wansink. 2015. "The impact of a supermarket nutrition rating system on purchases of nutritious and less nutritious foods." *Public health nutrition* 18:8-14.

Chen, C. 2007. "A finite smoothing algorithm for quantile regression." *Journal of Computational and Graphical Statistics* 16:136-164.

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.

Davis, G.C., and W. You. 2011. "Not enough money or not enough time to satisfy the Thrifty Food Plan? A cost difference approach for estimating a money–time threshold." *Food Policy* 36:101-107.

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

Devine, C.M., M.M. Connors, J. Sobal, and C.A. Bisogni. 2003. "Sandwiching it in: spillover of work onto food choices and family roles in low-and moderate-income urban households." *Social Science & Medicine* 56:617-630.

Devine, C.M., T.J. Farrell, C.E. Blake, M. Jastran, E. Wethington, and C.A. Bisogni. 2009. "Work conditions and the food choice coping strategies of employed parents." *Journal of Nutrition Education and Behavior* 41:365-370.

Drewnowski, A., and S. Specter. 2004. "Poverty and obesity: the role of energy density and energy costs." *The American journal of clinical nutrition* 79:6-16.

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.

Drichoutis, A.C., P. Lazaridis, and R.M. Nayga. 2005. "Nutrition knowledge and consumer use of nutritional food labels." *European Review of Agricultural Economics* 32:93-118.

Duan, N., W.G. Manning, C.N. Morris, and J.P. Newhouse. 1984. "Choosing between the sample-selection model and the multi-part model." *Journal of Business & Economic Statistics* 2:283-289.

Duffey, K.J., P. Gordon-Larsen, J.M. Shikany, D. Guilkey, D.R. Jacobs, and B.M. Popkin. 2010. "Food price and diet and health outcomes: 20 years of the CARDIA Study." *Archives of internal medicine* 170:420-426.

Eicher-Miller, H.A., V.L. Fulgoni, and D.R. Keast. 2012. "Contributions of processed foods to dietary intake in the US from 2003–2008: a report of the Food and Nutrition Science Solutions Joint Task Force of the Academy of Nutrition and Dietetics, American Society for Nutrition, Institute of Food Technologists, and International Food Information Council." *The Journal of nutrition* 142:2065S-2072S.

FDA. 2016. "FDA to Redefine “Healthy” Claim for Food Labeling." *Food and Drug Administration.* http://www.fda.gov/Food/NewsEvents/ConstituentUpdates/ucm520703.htm *(accessed December 6, 2016)*.

---. 2009. "Guidance for industry: letter regarding point of purchase food labeling." *Food and Drug Administration.* https://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm187208.htm *(accessed December 6, 2016)*.

Gagliardi, N. 2015. "Consumers Want Healthy Foods–And Will Pay More For Them." *Forbes. February* 18.

Gordon, E., N. Dawkins-Lyn, R. Hogan-Yarbro, A. Karpyn, K. Shore, S. Weiss, and S. Cash (2014) "Approaches for Promoting Healthy Food Purchases by SNAP Participants. Prepared by

ICF International for the U.S. Department of Agriculture, Food and Nutrition Service, July 2014. ." In.

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.

Greenwood, P.E., and M.S. Nikulin. 1996. *A guide to chi-squared testing*: John Wiley & Sons.

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

Grunert, K.G., and J.M. Wills. 2007. "A review of European research on consumer response to nutrition information on food labels." *Journal of public health* 15:385-399.

Gupta, S., P. Chintagunta, A. Kaul, and D.R. Wittink. 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.

Hao, L., and D.Q. Naiman. 2007. *Quantile regression*: Sage.

Harris, J.M., and R. Shiptsova. 2007. "Consumer demand for convenience foods: Demographics and expenditures." *Journal of Food Distribution Research* 38:22.

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* 71:1-14.

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.

Hunter, W., and T. Worsley. 2009. "Understanding the older food consumer. Present day behaviours and future expectations." *Appetite* 52:147-154.

IOM (2011) "Examination of Front-of-Package Nutrition Rating Systems and Symbols: Promoting Healthier Choices." In. Washington, DC The National Academies Press.

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)." *The American journal of clinical nutrition* 91:1102S-1108S.

Kennedy, P.E. 2005. "Oh no! I got the wrong sign! What should I do?" *The Journal of Economic Education* 36:77-92.

Kim, B.-D., and P.E. Rossi. 1994. "Purchase frequency, sample selection, and price sensitivity: The heavy-user bias." *Marketing Letters* 5:57-67.

Kim, H., L.A. House, G. Rampersaud, and Z. Gao. 2012. "Front-of-package nutritional labels and consumer beverage perceptions." *Applied Economic Perspectives and Policy*:pps037.

Koenker, R. 2005. *Quantile regression*: Cambridge university press.

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.

McCracken, V.A., and J.A. Brandt. 1987. "Household consumption of food-away-from-home: total expenditure and by type of food facility." *American Journal of Agricultural Economics*:274-284.

Morgan, K.J., M.E. Zabik, and G.L. Stampley. 1986. "Breakfast consumption patterns of US children and adolescents." *Nutrition Research* 6:635-646.

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.

Park, J.L., and O. Capps. 1997. "Demand for prepared meals by US households." *American Journal of Agricultural Economics* 79:814-824.

Prochaska, F.J., and R.A. Schrimper. 1973. "Opportunity cost of time and other socioeconomic effects on away-from-home food consumption." *American Journal of Agricultural Economics* 55:595-603.

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.

Reedy, J., and S.M. Krebs-Smith. 2010. "Dietary sources of energy, solid fats, and added sugars among children and adolescents in the United States." *Journal of the American Dietetic Association* 110:1477-1484.

Rose, D. 2007. "Food stamps, the Thrifty Food Plan, and meal preparation: the importance of the time dimension for US nutrition policy." *Journal of Nutrition Education and Behavior* 39:226-232.

Scott-Thomas, C. (2014) "WHO calls for standardised nutrition labelling. http://www.foodnavigator.com/Policy/WHO-calls-for-standardised-nutrition-labelling (accessed February 2, 2017)." In.

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.

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.

Strom, S. (2015) "TV Dinner in a Netflix World " In *The New York Times.* https://www.nytimes.com/2015/09/01/business/the-bustle-in-the-frozen-food-aisle.html (accessed February 12, 2017).

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.

Todd, J.E., L. Mancino, and B.-H. Lin. 2010. "The impact of food away from home on adult diet quality." *USDA-ERS Economic Research Report Paper*.

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.

Treiman, K., V. Freimuth, D. Damron, A. Lasswell, J. Anliker, S. Havas, P. Langenberg, and R. Feldman. 1996. "Attitudes and behaviors related to fruits and vegetables among low-income women in the WIC program." *Journal of Nutrition Education* 28:149-156.

Van Kleef, E., H. Van Trijp, F. Paeps, and L. Fernandez-Celemin. 2008. "Consumer preferences for front-of-pack calories labelling." *Public health nutrition* 11:203-213.

Variyam, J.N., J. Blaylock, and D. Smallwood. 1996. "A probit latent variable model of nutrition information and dietary fiber intake." *American Journal of Agricultural Economics* 78:628-639.

Wansink, B., and P. Chandon. 2006. "Can “low-fat” nutrition labels lead to obesity?" *Journal of Marketing Research* 43:605-617.

Wansink, B., and J. Sobal. 2007. "Mindless eating the 200 daily food decisions we overlook." *Environment and Behavior* 39:106-123.

Watson, E. (2015) "KIND petitions FDA to update ‘outdated’ food labeling rules." In. http://www.foodnavigator-usa.com/Regulation/KIND-petitions-FDA-over-healthy-nutrient-content-claim (accessed February 3, 2017).

--- (2017) "Should foods high in added sugar be disqualified from making nutrient content and health claims? ." In. http://www.foodnavigator-usa.com/Regulation/Union-of-Concerned-Scientists-files-added-sugar-citizen-petition (accessed February 3, 2017).

World Health Organization. 2015. "WHO calls on countries to reduce sugars intake among adults and children." *Retrieved February* 21:2016.

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., E.A. Finkelstein, J.M. Nonnemaker, S.A. Karns, and J.E. Todd. 2013. "Predicting the effects of sugar-sweetened beverage taxes on food and beverage demand in a large demand system." *American Journal of Agricultural Economics*:aat049.

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*

---. "Measuring the Informational Value of Interpretive Shelf Nutrition Labels to Shoppers." Agricultural and Applied Economics Association.

Zhu, C., R.A. Lopez, and X. Liu. 2015. "Information cost and consumer choices of healthy foods." *American Journal of Agricultural Economics*:aav057.

APPENDIX

A Log-Gamma Distribution for the second part of the TPM

As shown in figure 1.3, the distribution of the purchasing data appears to be skewed to the right. On average, 99.9% of the sample consists of zero purchases of UPC *i* during a shopping trip to store *r* made by household *h* at week *t* (Table 1.2). 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

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

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 $k = 0$ and an exponential distribution is captured by $\sigma = 1$.

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

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

B NuVal Effects Log-Gamma Distribution

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

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

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

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

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:

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

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

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

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

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

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

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

where *adopt* correspond to the NuVal variable.

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

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

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

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.

D Log-Normal Distribution for the second part of the TPM

The expected value of a conditionally lognormal variable is given by:

(28) $$E(y \mid y > 0; x) = \exp(x'\beta + \frac{1}{2}\sigma^2)$$

The dispersion parameter σ can be also estimated by dividing the Pearson goodness-of - fit statistic by its degrees of freedom. The Pearson's chi-squared statistic can be used for a test of a chosen distribution. Under the null hypothesis, the observed distribution of the outcome variable is consistent with a particular theoretical distribution (Greenwood and Nikulin, 1996).

E NuVal Effects Log-Normal Distribution

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

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

Where ϕ represents the standard normal cumulative distribution function or logit

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

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 NuVal effect TE^u can be defined as

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

Where *adopt* correspond to the NuVal variable.

The conditional NuVal effect TE^c can be computed as:

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

F Conditional 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 $E(y|x) = \phi(x'\alpha) * \exp(x'\beta + \frac{1}{2}\sigma^2)$.

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

(34) $e^c = \frac{dE(y|y > 0, x)}{dx_j} * \frac{x_j}{E(y|y > 0, x)} = \beta_j(\exp[x^{'}\beta + \ln(k^2)]) * \frac{x_j}{\exp[x^{'}\beta+\ln(k^2)]}$

(35) $e^c = \beta_j * x_j$

G 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 $E(y|x) = \phi(x'\alpha) * \exp(x'\beta + \frac{1}{2}\sigma^2)$.

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

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

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

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.

H Partial Effects of the TPM

	Whole Sample				Low Inc					High Inc				
	Participation Decision		Quantity Decision		Participation Decision			Quantity Decision		Participation Decision			Quantity Decision	
	Odds Ratio		Partial Effects		Odds Ratio			Partial Effects		Odds Ratio			Partial Effects	
P		***	-0.121	***	0.593	***		-0.149	***	0.531	***		-0.117	***
Adopt		**	-0.069		1.416	**		-0.723	***	0.682	***		-0.004	
Adopt*Score			0.007		0.980	***	-2.0%	0.023	**	1.012	***		0.008	
Ad		***	0.166	***	1.401	***		0.148	***	1.576	***		0.220	***
PR		***	0.074	*	**1.963**	***		0.047		1.808	***		0.117	**
NuVal Store		**	1.162	***	0.912	*		1.140	**	1.118	***		1.496	***
Households			1495					420					1075	
UPC			242					233					241	
QIC/AIC			19865		78555			6163		189283			14246	
N			19190		8316672			5709		21727601			13481	

	No College				College					Married				
	Participation Decision		Quantity Decision		Participation Decision			Quantity Decision		Participation Decision			Quantity Decision	
	Odds Ratio		Partial Effects		Odds Ratio			Partial Effects		Odds Ratio			Partial Effects	
P	0.550	***	-0.111	***	0.540	***		-0.091	***	0.630	***		-0.025	
Adopt	0.956		-0.111		0.766	**		-0.072		1.398			-0.020	
Adopt*Score	1.000		0.010	*	1.005			0.007		0.976	*	-2.4%	0.003	
Ad	1.582	***	0.056		1.465	***		0.161	***	**1.792**	***		0.138	***
PR	1.831	***	0.065	**	1.851	***		0.045		1.822	***		0.101	***
NuVal Store	1.013		0.579	*	1.106	***		0.410		1.092			-0.138	
Households			859					636					172	
UPC			236					239					203	
QIC/AIC	156383		11923		111233			8511		32278			2434	
N	17319679		11291		12724594			7899		3533873			2342	

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

I Partial Effects of the TPM (Continued)

	Single				Children				No Children			
	Participation Decision		Quantity Decision		Participation Decision		Quantity Decision		Participation Decision		Quantity Decision	
	Odds Ratio		Partial Effects		Odds Ratio		Partial Effects		Odds Ratio		Partial Effects	
P	0.549	***	-0.081	**	0.560	***	-0.117	**		***	-0.125	***
Adopt	0.754	***	-0.099		0.647	***	-0.450	*			0.246	**
Adopt*Score	1.007		0.009	*	1.021	***	0.023	**			-0.006	
Ad	1.546	***	0.167	***	1.500	***	0.167	*		***	0.176	***
PR	1.937	***	0.085	**	1.794	***	-0.004			***	0.120	***
NuVal Store	1.095	***	0.815	***	1.212	***	1.465	***			1.145	***
Households			986				301				1194	
UPC			240				230				241	
QIC/AIC	171464		12787		56308		4252		211633		15929	
N	20587344		12139		5563102		3993		24481171		15197	

	Full Time				Retired				Not employed			
	Participation Decision		Quantity Decision		Participation Decision		Quantity Decision		Participation Decision		Quantity Decision	
	Odds Ratio		Partial Effects		Odds Ratio		Partial Effects		Odds Ratio		Partial Effects	
P	0.628	***	-0.091	*	0.658	***	-0.032	***	0.561	***	-0.138	***
Adopt	0.663	*	-0.218		1.092		-0.144		1.850	**	0.120	
Adopt*Score	1.013		0.015		0.991		0.006		0.966	*** -3.4%	-0.010	
Ad	1.768	***	0.129	*	1.561	***	0.071	**	1.461	***	-0.018	
PR	1.898	***	0.155	***	1.779	***	0.057		1.820	***	0.133	***
NuVal Store	1.034	*	0.781		1.016	***	0.260		1.138		0.523	
Households			330				408				153	
UPC			214				229				215	
QIC/AIC	54525		4169		80555		6414		28027		1964	
N	7642444		3907		7385489		5880		3153017		2031	

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