Information Cost and Consumer Choices of Healthy Foods

Chen Zhu, Rigoberto A. Lopez, and Xiaoou Liu*

This article examines whether or not a reduction in consumer search cost for nutritional information increases the probability that heterogeneous consumers will choose healthier food products. Empirical results from the ready-to-eat breakfast cereal (RTEC) market confirm the conceptual analysis that lowering information cost via simplified nutritional labeling increases the healthfulness of consumer choices. The healthfulness attribute weighs 28.44% more heavily in consumers' decision-making with simpler labeling systems. On average, introducing front-of-package labeling increased the probability of a consumer choosing a healthy RTEC by 3.49% and reduced the probability of choosing an unhealthy RTEC by 3.81%. Calories, sugar, saturated fat, and sodium consumption decrease by 0.31%, 2.63%, 6.94%, and 1.97%, respectively. Fiber intake increases by 3.24%. Further results show that less-educated and smaller households with less frequent purchases benefit the most from a reduction in information cost. Overall, this article shows the potentially positive role that voluntary, more convenient labeling could play in improving market and public health outcomes.

Key words: Healthfulness, information cost, consumer choices, food labeling.

JEL codes: D12, I19, L66, M30.

Unhealthy consumer food choices remain at the center of the public health battle against obesity and other food-related diseases. By providing information, nutrition labels provide a means for encouraging healthier food choices. Based on this rationale, the US government implemented its first mandatory labeling system, the Nutrition Facts Panel (NFP), under the nutrition Labeling and Education Act of 1990 (NLEA). The NFP labeling system aims to enhance public health by requiring food companies to print detailed nutritional information on both quantities per serving and percentage of the daily value on the back or side of food packages.

Empirical evidence shows, however, that NFP labeling has not been effective at improving the healthfulness of consumer food choices. The provision of information alone does not guarantee its use. Many consumers do not understand NFP labels (Rothman et al. 2006; Visschers, Hess, and Siegrist 2010). Only 53% of consumers report ever using NFP information, and the usage has been declining (Blitstein and Evans 2006; Todd and Variyam 2008). In fact, the NFP causes no change in consumers' search behaviors (Balasubramanian and Cole 2002). Recent evidence even reveals that the NFP has been counterproductive in terms of promoting healthfulness of food choices—an unintended consequence of the regulation (Moorman, Ferraro, and Huber 2012; Wang, Rojas, and Bauner 2015).[1]

*Chen Zhu is an assistant professor at the College of Economics and Management, China Agricultural University in Beijing, China. Rigoberto Lopez is a professor at the Department of Agricultural and Resource Economics, University of Connecticut in Storrs, CT. Xiaoou Liu is an associate professor at the School of Agricultural Economics and Rural Development, Renmin University of China in Beijing. The research is funded by the National Science Foundation of China (Project No. 71103188 & 71373268), the National Social Science Fund of China (14CJY018), the Chinese Universities Scientific Fund (2015QC010), China Postdoctoral Science Foundation Grant (2013M542390 &2014M560148), the Zwick Center for Food and Resource Policy at the University of Connecticut, and the Center for Food and Health Economic Research at China Agricultural University. The authors thank *AJAE* editor Madhu Kanna and two anonymous reviewers for their constructive comments. Correspondence may be sent to: xiaoou.liu@ruc.edu.cn.

[1] Moorman, Ferraro, and Huber (2012) point out that the NLEA reduced the brand nutritional quality relative to a control group, while Wang, Rojas, and Bauner (2015) find that the nutrition quality of ready-to-eat breakfast cereals decreased between 1988 and 2001.

Amer. J. Agr. Econ. 98(1): 41–53; doi: 10.1093/ajae/aav057
Published online November 10, 2015

An explanation for the ineffectiveness of the NFP is that there is a high information cost when consumers have to process complicated nutritional information (Kiesel and Villas-Boas 2013; Levy and Fein 1998; Berning, Chouinard, and McCluskey 2008).[2] Instead of the NFP, consumers prefer simpler labels with summarized key nutritional facts printed conveniently on the front of a package (Williams 2005; Wansink, Sonka, and Hasler 2004; Grunert and Wills 2007), and they are more likely to use information in simplified formats, particularly for calories (Bollinger, Leslie, and Sorensen 2011).[3] Kiesel and Villas-Boas (2013) find that summarized and simple nutritional labels have a positive impact on sales of microwave popcorn but that the impact diminishes when labels become complicated.[4]

Since 2007, some food companies have voluntarily introduced simplified nutritional label systems such as a Front of Package Panel (FOP) summarizing nutritional information for four key nutrients.[5] This significantly reduces consumers' information cost without providing new nutritional information, as the NFP is still printed on the back of the package (Zhu and Huang 2014). This change of information provision formats in real markets provides a unique opportunity to test if information cost reduction increases the probability that consumers will choose healthier foods.

This article extends the food labeling literature in several ways. First, most previous studies focus on consumer label use or the impact of nutritional labeling on market sales. This article focuses on the impacts of introducing a simplified labeling system *in addition* to the pre-existing labeling system (NFP) on the healthfulness of choices by deriving a direct link between the information cost imposed by different labeling systems and the healthfulness of consumer choices. Second, rather than relying on controlled experiments or surveys (e.g., Variyam 2008; Kiesel and Villas-Boas 2013; Liaukonyte et al. 2013), this article analyzes a rich dataset containing approximately 128,629 actual household-level brand purchases matched to brand-level nutritional information, advertising, and labeling.[6] We apply a random coefficient mixed logit model to consumer ready-to-eat breakfast cereal (RTEC) choices, allowing for labeling formats and consumer heterogeneity. We then multiply the 128,629 real purchases by 22 brands and generate 2,829,838 observations to match in the mixed logit model.

Empirical results from the U.S. RTEC market indicate that, on average, FOP increases the probability of consumers choosing healthier foods, and decreases the probability that they will choose a less healthy option. As a consequence, consumers decrease their consumption of sugar, saturated fats, and sodium (designated as "bad" nutrients), but increase their consumption of fiber (designated as a "good" nutrient). We also find that less-educated consumers from small households who purchase less frequently are more sensitive to information cost changes. This finding implies that simpler labels have a significantly larger impact on less-educated consumers and those from smaller households who purchase less frequently, as they may not have been as aware of the healthfulness information on the NFP before the introduction of FOP.

Conceptual Analysis

To link information cost to the healthfulness of consumers' food choices, we extend

[2] There are several alternative explanations. For example, many consumers knowingly consume high-calorie and unhealthy food because of time constraints, a limited food environment, or lack of food preparation skills (Gregory, Rahkovsky, and Anekwe 2014). Consumers may limit their attention or ration information depending on their perceived benefits v. their marginal search costs, and optimally allocate their limited attention to information (Matejka and McKay 2015). Additionally, certain consumers' lack of interest in nutrition quality may be due to their negative inferences about the taste of highly nutritional options (Moorman, Ferraro, and Huber 2012; Chidmi and Lopez 2007; Van Wezemael et al. 2014). In this article, we focus on the explanation that people ignore information because of high information cost (e.g., complexity).

[3] Many restaurants and fast food chains are expected to post nutritional information, such as calories, as required by the 2010 Patient Protection and Affordable Care Act.

[4] Like Wansink and Chandon (2006), Kiesel and Villas-Boas (2013) find that the effect of labeling products for "low fat" has a net negative effect on consumer choices and, in their study, a negative effect on supermarket sales of popcorn.

[5] In 2012, the FOP system became known as the "Facts Up Front" system.

[6] This section assumes that ready-to-eat breakfast cereal purchase decisions are single-person decisions made by one household head rather than a group decision that incorporates preferences from everyone in a household.

consumer search theory to include "healthfulness" as a specific product attribute, and use models of consumer search for best-matched product alternatives with sequential learning (Weitzman 1979; Wolinsky 1986; Bakos 1997; Anderson and Renault 1999; Silberberg and Suen 2000; Armstrong, Vickers, and Zhou 2009).

Let $v_i(p_j, h_j, \boldsymbol{Z}_j, \boldsymbol{D}_i)$ denote the indirect utility function of consumer i when purchasing a certain brand j in a market; p_j and h_j are the price and healthfulness of brand j, $\boldsymbol{Z}_j$ is a vector that collects other product attributes, and $\boldsymbol{D}_i$ is a vector of consumer demographic characteristics that affect utility. Without loss of generality, we simplify $v(\cdot)$ by dropping subscript i and variables $\boldsymbol{D}_i$, p_j, and $\boldsymbol{Z}_j$ in the following discussion in order to focus on the healthfulness attribute of the product.[7]

In a market with imperfect information, the consumer knows the healthfulness distribution over all brands in his/her choice set $f_h(h)$, but *ex ante* not the exact value of healthfulness of each brand j. That is, the consumer is aware of healthfulness value levels available in his/her choice set without knowing which brand has a particular value of h. The range of h is $[a, b]$, where a and b give the lowest and highest healthfulness measurements, respectively, of products available in the consumer's choice set.

The consumer searches using a sequential search strategy. After reading a brand's nutritional information, he/she evaluates whether to continue to search or stop and accept the current brand. The consumer stops searching and finalizes the purchase when the information cost becomes too high to proceed to another search. Let a consumer start the search from brand k with healthfulness h_k. His/her expected gain from an additional search, $G(h_k)$, is defined as:

$$(1)\qquad \begin{aligned} G(h_k) &= E(gains \mid h_k) \\ &= E(v(h_{k+1}) - v(h_k) \mid h_k) \\ &= \int_{h_k}^{b} (v(h_{k+1}) - v(h_k)) f(h_{k+1}) dh_{k+1}. \end{aligned}$$

Integrating yields

$$(2)\qquad \begin{aligned} G(h_k) = b - v(h_k) \\ - \int_{h_k}^{b} F_h(h_{k+1}) v'(h_{k+1}) dh_{k+1}. \end{aligned}$$

As shown in the online supplementary appendix 1, $G(h_k)$ is a decreasing function of h_k when the marginal utility of healthfulness is greater than 0. This property ensures a diminishing marginal return to $G(h_k)$ on searching. The higher the healthfulness of the current brand is, the lower are the expected gains from further search. When h_k is sufficiently high, the expected gain from an additional search falls below the marginal information cost c. Thus, the reservation level of healthfulness, h^*, is defined by the equilibrium condition:

$$(3)\qquad \begin{aligned} G(h^*) = b - v(h^*) \\ - \int_{h^*}^{b} F_h(h_{k+1}) v'(h_{k+1}) dh_{k+1} = c. \end{aligned}$$

Equation (3) implies that if h_k is below the reservation healthfulness level h^*, the search continues. Otherwise the consumer stops searching.[8]

Applying the implicit function theorem to equation (3), we can solve h^* as a function of c, that is, $h^* = h^*(c)$. Substituting h^* back into equation (3) and taking the derivative with respect to c, we obtain:

$$(4)\qquad \frac{\partial h^*}{\partial c} = \frac{1}{v'(h^*)[F(h^*) - 1)]} < 0.$$

Thus, lowering the information search cost increases the reservation healthfulness level h^* for a given consumer.[9]

The probability that the consumer will continue searching is:

$$(5)\qquad \begin{aligned} &\mathrm{Prob}(continue\ to\ search) \\ &\quad = \mathrm{Prob}(v(h_k) \le v(h^*) - c) = F_h(v(h^*) - c). \end{aligned}$$

[7] In a realized purchase, a consumer makes purchase decisions based on all product attributes, such as price, healthfulness, taste, color, and convenience, and demographic variables like income, etc. We control for other variables in the empirical analysis, but in the conceptual model we focus solely on healthfulness to evaluate how information cost changes affect a consumer's concern for the healthfulness attribute of a product.

[8] The marginal search cost may vary across consumers. Consumers differ in their information-gathering costs due to differences in analytic ability, the cost of time, and preference for reading and processing information. In our empirical analysis, we assume the change in information cost is identical for all consumers.

[9] Details are in the supplementary appendix 1 online.

A higher h^* indicates that the consumer's probability of continuing to search is higher; consumers search more intensively and the probability of them selecting a healthier brand is higher.[10]

We summarize the implications from equations (4) and (5) as follows: decreasing information search cost c increases a consumer's reservation healthfulness level h^*, which in turn increases the probability that a consumer will choose a healthier product. According to equation (4), for heterogeneous consumers i and n who preserve the same reservation level of h^*, their responses to information cost are different when they have different marginal utilities of healthfulness. The consumer who has a smaller marginal utility of healthfulness is more sensitive to information cost changes. That is, consumers who care less about healthfulness will benefit more from information cost reduction.

Empirical Framework

The empirical analysis tests whether a decrease in information cost has a positive impact on the probability that heterogeneous consumers will choose healthier RTECs. The introduction of FOP is a proxy variable for the reduction of information cost. The RTEC market provides a good case study for several reasons. First, it offers a natural experiment for the empirical test. The biggest two RTEC manufacturers, Kellogg's and General Mills, adopted a FOP labeling system for selected brands in October 2007, while their competitors, Quaker Oats and Post, continued to display only the NFP on the back of the package for all brands.[11] Second, RTEC products provide a wide array of choices with respect to nutritional content and healthfulness, ensuring significant variation in the data. Consumer choices in the RTEC market have been extensively studied in previous work (For example, Kiser 1998; Nevo 2001; Chidmi and Lopez 2007), thus allowing consumer choice data to be for compared and validated within the context of labels and healthy RTEC choices.

Mixed Logit Model for Consumer Response to Labeling

To analyze the data we use a mixed logit model, a highly flexible procedure that removes the well-known Independence of Irrelevant Alternatives property of the standard logit model and allows for random taste variation across consumers. This model incorporates individual-level information efficiently and is less computationally intensive compared to other random coefficient models.[12]

Let consumer i choose an RTEC brand j from among competing products $\{1,\ldots,J\}$ in market m to maximize the conditional indirect utility:

$$(6) \qquad v_{ijm} = \alpha_i p_{jm} + \boldsymbol{Z}'_{jm}\beta_i + h_j(\gamma_{1i} + \gamma_{2i}FOP_j) + \xi_{jm} + \varepsilon_{ijm}.^{13}$$

Variable p_{jm} is the price of RTEC brand j, h_j is the healthfulness of RTEC j, FOP_j is a proxy variable for information cost, $FOP_j = 1$ represents RTEC product j displaying FOP and thereby a decrease in information cost, $FOP_j = 0$ is a benchmark impact of healthfulness on consumers without changes in information cost, $\boldsymbol{Z}_{jm}$ denotes other product attributes of j, ξ_{jm} is unobserved product characteristics that may correlate with price, and ε_{ijm} is an independent and identically distributed disturbance term.[14] Coefficients α_i, β_i, γ_{1i}, γ_{2i} are randomized across heterogeneous consumers.

To capture different responses to FOP by heterogeneous consumers, we further specify the random coefficient γ_{2i} to be composed of

[10] Although consumers' reservation healthfulness level h^* increases with a reduction of the information cost, the final consumer choice will also depend on the supply side in the market. It is likely that if there is no brand with healthfulness close to h^*, consumers will choose the same brand as before the information cost change. We can only conclude that when consumers search more intensively, their probability of making a healthier choice increases.

[11] The dates for FOP printed on packages are listed in table 1.

[12] For instance, using the Berry, Levinsohn, and Pakes' (1995) random coefficients logit model would allow for consumer heterogeneity, but aggregating responses to a market level and using a sample of consumers for each market (instead of observations on all consumers) would result in a loss of efficiency.

[13] Here, "conditional" means conditioned on a non-purchase option. The mixed logit model is not designed to include a non-purchase option because no data on outside market product characteristics are available. The utility specified in mixed logit is considered to be a conditional utility, and the individual choice probability is a conditional probability (Train 2009).

[14] The explanatory variable, FOP, may correlate with the variance of the error term and cause heteroskedasticity. The mixed logit model may correct for heteroskedasticity when it is interpreted as an error component model by collecting the stochastic portion of the random coefficients together with the iid extreme value (Train 2009).

fixed and variable components that change with consumers' observed demographics $\boldsymbol{D}_i$ and unobserved characteristics ζ_i,

$$\text{(7)} \qquad \gamma_{2i} = \bar{\gamma}_2 + \lambda \boldsymbol{D}_i + \kappa \zeta_i$$

where λ is a matrix with elements measuring how consumers' preferences vary with demographics, and ζ_i follows standard multivariate normal distribution with κ as a scaling vector.

Price endogeneity is a less important problem for an individual-level analysis because price is exogenous to consumers. But if ξ_{jm} is not separated from the disturbance term, endogeneity arises when the unobserved product attributes ξ_{jm} correlate with price. To deal with the unobserved variable ξ_{jm}, we use a control function approach proposed by Petrin and Train (2010). We define a control function $CF(\cdot)$ as

$$\text{(8)} \qquad \xi_{jm} = CF(\eta_{jm}; \phi)$$

where η_{jm} is the residual from a regression of the price variable on the observed explanatory and instrumental variables, and ϕ is the coefficient corresponding to η_{jm}. We then estimate the mixed logit model with the control function entering as an extra variable. We specify a linear function as the control function $\xi_{jm} = CF(\eta_{jm}; \phi) = \phi\eta_{jm}$ for simplicity. Substituting equations (7) and (8) into (6), the final utility specification is as follows:

$$\text{(9)} \qquad \begin{aligned} v_{ijm} = {} & \alpha_i p_{jm} + \boldsymbol{Z}'_{jm}\beta_i + h_j\gamma_{1i} \\ & + (\bar{\gamma}_2 +_j \lambda \boldsymbol{D}_i + \kappa\zeta_i) h_j FOP_j \\ & + \phi\eta_{jm} + \varepsilon_{ijm}. \end{aligned}$$

When consumers purchase a unit of a brand, the probability that consumer i purchases a unit of brand j in market m is

$$\text{(10)} \qquad P_{ijm} = \int_{\zeta} \frac{\exp(V_{ijm})}{\sum_{r=1}^{J} \exp(V_{irm})} g(\zeta) d\zeta.$$

Impact of FOP on Consumer Healthy Choices

The coefficient of the interaction term $h_j FOP_j$ captures how a consumer's choice responds to information cost reductions. Changes in the marginal effect of healthfulness on choice probability $\partial P_{ijm}/\partial h_j$ due to the adoption of FOP are equal to:

$$\text{(11)} \qquad \frac{\Delta(\partial P_{ijm}/\partial h_j)}{\Delta FOP_j} = \int_{\zeta} (\bar{\gamma}_2 + \lambda \boldsymbol{D}_i) L_{ijm}(\zeta)(1 - L_{ijm}(\zeta)) g(\zeta) d\zeta$$

where $L_{ijm}(\zeta)$ is the conditional probability on ζ.[15] If estimates of $\frac{\Delta(\partial P_{ijm}/\partial h_j)}{\Delta FOP_j}$ take positive signs, the healthfulness attribute of a product will weigh more in consumers' purchase decisions because of the introduction of FOP.

The simulated probability for consumer i to choose brand j with and without FOP is

$$\text{(12)} \qquad \breve{P}^{w}_{ijm} = \frac{1}{R}\sum_{r=1}^{R} \hat{L}^{w}_{ijm}(\tilde{\zeta}^r) \quad \text{and}$$

$$\breve{P}^{wo}_{ijm} = \frac{1}{R}\sum_{r=1}^{R} \hat{L}^{wo}_{ijm}(\tilde{\zeta}^r)$$

where $\tilde{\zeta}^r$ is the rth draw from $g(\zeta)$, and $\hat{L}(\cdot)$ is the estimated conditional probability. If the differences between the two simulated probabilities increase with healthfulness, we can conclude that FOP increases consumers' probability of choosing healthier foods.

Data

The data used to operationalize the model are taken from multiple sources. After matching these sources, purchase behaviors for 5,844 households over 152 weeks are used for our estimation, along with three alternative measures of RTEC healthfulness at the brand level. The resulting real purchases in the sample are 128,629. We multiply 128,629 real purchases by 22 brands and generate 2,829,838 observations to match in the mixed logit model.

Measures of Brand-level Healthfulness

The healthfulness of foods is a multidimensional concept for which there are no absolute measures. This concept can

[15] Details of the derivations are in the online supplementary appendix 2.

be broken down into the nutritional content of foods, as recommended by public agencies like the USDA or the Institute of Medicine, which provide advice on what nutrients Americans are consuming in excess amounts and which ones should be avoided to improve one's health.[16] To check for robustness of the analysis, we use three alternative measures at the RTEC brand level: (1) the Nutrition Profile Index; (2) NuVal scores, and (3) indexes based on principal component analysis of nutrients in RTEC brands.

The Nutrition Profile Index (NPI) scores reflect food quality assessments and are calculated based on a model developed for the Food Standards Agency (FSA) of the United Kingdom (Castetbon, Harris, and Schwartz 2011). Rather than relying on a single nutrient measurement, the NPI scores account for both positive (e.g., protein, fiber, vitamins) and negative (e.g., sugar, sodium, saturated fat) nutrients in the entire nutrient composition, providing a comprehensive evaluation of the nutritional quality of food products.

NuVal LLC scores food products from 1 to 100 based on the Overall Nutritional Quality Index (ONQI) algorithm (Zhen and Zheng 2015). Products with higher scores are considered by the ONQI expert panel to be healthier than products receiving lower scores. The algorithm places health-favorable and health-unfavorable nutrients in the numerator and the denominator, respectively. The impacts of other key nutrition factors that measure the quality and density of nutrients as well as the strength of their association with specific health conditions are also included. The NuVal scoring system represents validated expert opinion on the healthfulness of a food product. One should note that NuVal scores are displayed on shelf price tags in NuVal LLC-collaborating retail stores, not on packages, and that their introduction in selected stores happened later than the RTEC sample used in this article.[17]

Principal component analysis (PCA), an exploratory multivariate statistical technique for summarizing data, is discussed in the online supplementary appendix 3. The selection of the PCA indicators is the result of carefully screening key nutrients that determine the healthfulness of USDA recommendations, and those of private providers of healthfulness information such as NPI and NuVal. To this end, we identified five nutritional characteristics: calories (which include not only sugar but all sources), sugar content, sodium, saturated fats, and fiber (the only "good" nutrient included). We collected information on the nutrient characteristics of the RTECs in our sample and augmented the sample to include some obviously "healthy" foods with NuVal scores equal to 100 (e.g., tomatoes, lettuce, broccoli, apples) and some obviously "unhealthy" foods (carbonated soft drinks at the brand level, selected salty snacks, and candy).

Table 1 lists the three healthfulness measurements for all brands used in the analysis. The NPI scores of brands in the sample range from 26 to 74, NuVal from 4 to 33, and PCA from 6 to 69. The correlation coefficients between the PCA index and NPI and NuVal scores are 0.8551 and 0.8155; they are higher than the correlation between NPI and NuVal scores, which is 0.7932.

Data on Household Purchases, Demographics, Advertising, and Labeling

These data are from the following three proprietary datasets covering January 2006 to December 2008: RTEC household purchases data from Nielsen Homescan; RTEC product-level weekly TV advertising exposure from Nielsen Media Research; RTEC packaging information, including labeling and nutritional information, from Mintel Global New Products Database (GNPD).

The Nielsen Homescan database tracks purchases of RTECs across Designated Market Areas (DMAs) in the United States.[18] The purchases for at-home consumption

[16] For instance, in terms of "negative" nutrients, Americans have been advised to consume less sugar, sodium, and saturated fats. On the "positive" nutrient side are "low fat" and fiber content (e.g., whole instead of refined grains). Volpe, Okrent, and Leibtag (2013) provide several methods for measuring healthfulness, based on USDA recommendations, of a consumer basket of goods at various retail outlets for aggregate categories of food. Our measures are at the brand-level of foods, but they are also driven by nutrient content, weighted according to what is recommended as healthy by the USDA, as well as NuVal and NPI scores.

[17] We collected data on NuVal scores from local stores. The NuVal scores for all brands owned by Kellogg's, General Mills, Quaker and two brands owned by Post (Grape-Nuts and Honey Bunches of Oats) were collected from Price Chopper in Storrs, Connecticut. The NuVal scores for two other brands owned by Post (Fruity Pebbles and Cocoa Pebbles) were collected from Big Y in Mansfield, Connecticut.

[18] Our database includes the following 16 DMAs: New York, Philadelphia, Detroit, Boston, Washington DC, Baltimore, Atlanta, Miami-Ft. Lauderdale, Hartford-New Haven, Springfield-Holyoke, Chicago, Kansas City, Houston, Los Angeles, San Francisco-Oakland-Santa Rosa, and Seattle-Tacoma-Bellingham.

Table 1. Healthfulness Measurements of Top Ready-to-eat Cereals

Firm	Brand	NPI	NuVal	PCA	Date for FOP printed on the package
Kellogg's	Frosted Flakes	42	11	58	08-8-13
Kellogg's	Frosted Mini-Wheats	74	33	69	08-3-11
Kellogg's	Raisin Bran	54	27	65	08-3-13
Kellogg's	Froot Loops	39	23	37	08-4-22
Kellogg's	Rice Krispies	41	22	57	08-3-13
Kellogg's	Special K	44	23	59	08-2-13
Kellogg's	Special K Red Berries	48	23	57	08-3-5
Kellogg's	Apple Jacks	40	11	54	08-8-22
Kellogg's	Corn Pops	33	20	55	07-12-7
General Mills	Cheerios	58	33	66	08-3-12
General Mills	Cinnamon Toast Crunch	37	21	51	08-3-19
General Mills	Lucky Charms	36	23	52	08-6-19
General Mills	Cocoa Puffs	39	23	53	08-7-17
General Mills	Reese's Puffs	34	23	50	08-7-21
Quaker	Cap'n Crunch	28	4	34	No FOP
Quaker	Life Cinnamon	53	24	61	No FOP
Quaker	Cap'n Crunch Crunchberries	28	4	34	No FOP
Quaker	Cap'n Crunch Peanut Butter Crunch	32	8	36	No FOP
Post	Honey Bunches of Oats	54	24	59	No FOP
Post	Grape-Nuts	70	28	67	No FOP
Post	Fruity Pebbles	26	8	38	No FOP
Post	Cocoa Pebbles	26	7	39	No FOP

include buying at big box retailers, grocery stores, convenience stores, automatic vending machines, and online retailers.[19] This dataset is the main resource for explanatory variables used in the analysis.

The Nielsen Media Research database provides brand-level TV advertising exposure on a weekly basis for each DMA, measured in gross rating points (GRPs). A higher GRP means more consumers were exposed to a brand's TV advertising aired in a given area and week. As the most heavily advertised food product category in the United States, RTEC advertising is a key non-price variable affecting choices in the relevant markets.

The Mintel GNPD has provided detailed product listings for 245 categories of food, drink, and other grocery store items since 1996. Product listings are collected by Mintel based on product reformulations, new product introductions, new product packaging, and new product varieties. This dataset is used to collect precise introduction of FOP label information for each brand in the sample.

Table 2 reports the summary statistics for all the variables used in the analysis. Although highly correlated, NuVal offered the lowest mean score of healthfulness, while PCA offered the highest. Only about 15% of the observations had an FOP label, which represents a rather modest share. The average household size was approximately three persons, with 23% of the household heads being college educated. Average frequency of a household purchasing RTEC is about 31 weeks in the 152-week period. Finally, 56.1% of households earn more than \$60,000 per year, and 93.9% of the primary shoppers in the sample are female.[20]

We did not select these DMAs; rather, they constitute the full dataset purchased by our data provider.

[19] For each purchase, the dataset reports time and location of the purchase, price and quantity, product characteristics such as brand and package size, and demographic characteristics of buyers.

[20] Although these data sources are proprietary, under restricted contractual conditions, researchers may access the data used in this study solely for purposes of reproduction. For purposes beyond reproduction, researchers are directed to access these data by directly contacting the Nielsen and Mintel companies. The healthfulness indexes used are found in table 1. The Stata computer programming codes for the mixed logit estimation are available from the authors upon request.

Table 2. Summary Statistics for Explanatory Variables

Variable	Definition	Mean	Std. Dev.	Min.	Max.
Product characteristics					
price	Dollars per ounce	0.164	0.057	0.000	0.500
Healthfulness: NPI	Nutrition Profile Index scores	43	13	26	74
Healthfulness: NuVal	NuVal scores	19	9	4	33
Healthfulness: PCA	Healthfulness scores calculated by PCA	52	11	34	69
Advertising	TV Advertising GRP	0.933	0.922	0	5.082
FOP	Dummy variable; FOP = 1 with FOP label printed	0.151	0.358	0	1
General Mills	Firm dummy variable	0.227	0.419	0	1
Quaker	Firm dummy variable	0.182	0.386	0	1
Post	Firm dummy variable	0.182	0.386	0	1
Last purchase dummy	Last purchase dummy = 1 if consumer purchase the same brand last time	0.024	0.153	0	1
Consumer demographics					
HHsize	Household size	3.120	1.472	1	9
HHedu	Dummy variable; HHedu = 1 if none of household heads has college degree or above	0.231	0.422	0	1
HighIncDum	Dummy variable; HighIncDum = 1 if average family annual household income is more than $60,000	0.561	0.496	0	1
women	Dummy variable; women = 1 if the household primary shopper is female	0.939	0.240	0	1
active	Household frequency of RTEC purchases	30.966	19.980	1	132

Empirical Results

In the estimation, we allow coefficients of healthfulness (h_j) and the interaction term between healthfulness and FOP ($h_j FOP_j$) to vary with consumer demographics. Coefficients of price (p_{jm}) and other product attributes (Z_{jm}) are estimated as following normal distributions, as in Train (2009). The instrumental variables used in the estimations are cost shifters, including firms' advertising expenditures, price of sugar, and price of wheat (Nevo 2001; Berry, Levinsohn, and Pakes 1995). The regression of price on all explanatory variables and instrumental variables for the control function has statistically significant coefficients for all instruments used. The high value of F statistics also implies that the instruments are strong.

Table 3 presents the selected estimates for equation (9) of the mixed logit model for healthfulness measurement NPI.[21] The estimated coefficient of price is −5.175 and significant at the 1% confidence level. Estimates of own price elasticities range from −2.856 to −1.540, which are comparable to Nevo (2001) and Chidmi and Lopez (2007) for brand-level RTECs. The estimate of TV advertising is 0.3212 and implies a positive relationship between choice probability and advertising. All estimated coefficients about firm dummies take a negative sign and indicate that Kellogg's brands are more

[21] The complete estimated coefficients for the NuVal and PCA healthfulness measurements are available in the online supplementary appendix 4. The coefficients estimates take the same signs as the main variables. The calculated impact and probability changes are also similar to NPI's.

Table 3. Estimates of the Mixed Logit Model for RTEC Choices

Variable	Mean	Std. Err.	Std. Dev.	Std. Err.
price	−5.1752***	(0.0706)	5.4509***	(0.0773)
advertising	0.3212***	(0.0054)	0.3119***	(0.0076)
Last purchase dummy	1.9556***	(0.0132)	0.7372***	(0.0219)
NPI	0.0377***	(0.0005)	0.0470***	(0.0005)
FOP × *NPI*	0.0107***	(0.0018)	0.0004	(0.0004)
FOP × NPI × *HHsize*	−0.0051***	(0.0006)		
FOP × NPI × *HHedu*	0.0013**	(0.0006)		
FOP × NPI × *HighIncDum*	0.0005	(0.0006)		
FOP × NPI × *women*	0.0016	(0.0011)		
FOP × NPI × *active*	−0.0014***	(0.0005)		
General Mills	−0.1753***	(0.0158)	1.4579***	(0.0168)
Quaker	−1.0834***	(0.0222)	1.2947***	(0.0218)
Post	−0.8289***	(0.0178)	1.5790***	(0.0194)
Control function residuals η	63.9033***	(0.3720)		
Observations		2,829,838		
Log Likelihood		−297905.284		
First step F statistics		17514.24		
p-value		0.0000		

Note: Standard errors appear in parentheses. Asterisks *** represent a p-value $p < 0.01$; ** represents a p-value $p < 0.05$; *represents a p-value $p < 0.1$. Kellogg is used as the benchmark of firm dummies.

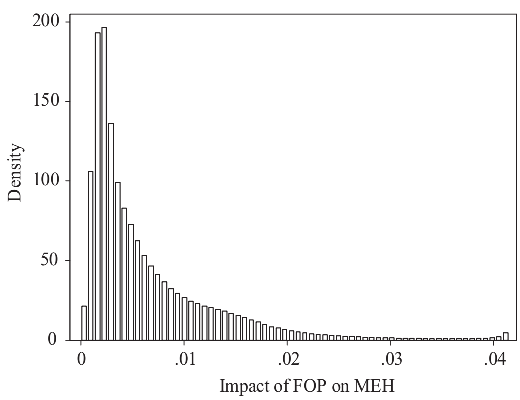

Figure 1. Density distribution for the impact of FOP on marginal effect of healthfulness

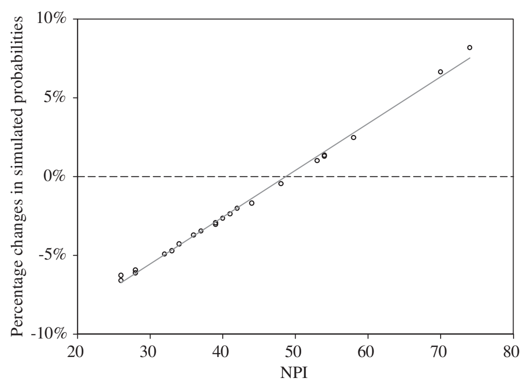

Figure 2. Changes in simulated probability with and without FOP

likely to be chosen. The estimate of the last purchase dummy indicates that last purchase behaviors are positively related to consumers' choice probabilities. Female shoppers and high-income households are indifferent in term of responsiveness to FOP compared with male shoppers and low-income households because the estimated coefficients on interaction terms of women and HighIncDum are not statistically significant.

The impact of FOP on the marginal effect of healthfulness in equation (11) is calculated for each observation. The distribution of this impact is shown in figure 1. The positive sign for all observations indicates that FOP has a positive and significant impact on the marginal effect of healthfulness. The calculated average impact is 0.0086 and implies that the healthfulness attribute of RETC products weighs, on average, 28.44% more heavily in consumers' purchase decisions when FOP is displayed on the package.

The scattered points in figure 2 show the percentage changes in simulated probabilities for each brand being chosen with and without FOP. For example, in our sample, the simulated probability for consumers to purchase a brand with NPI = 74 is 13.41% without FOP. The probability increases by 8.31% to 14.95% with the introduction of

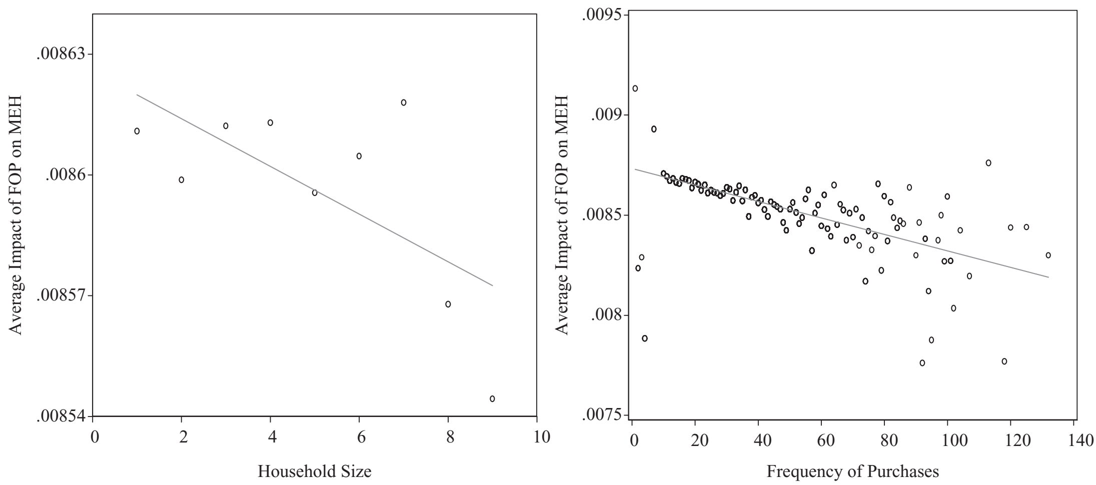

Figure 3. Average FOP impact on marginal effect of healthfulness by household sizes and frequency of purchases

FOP. For a product with lower NPI, say, NPI = 26, the simulated probability that consumers will choose it decreases by 7.74% from 1.55 % to 1.43% because of FOP. The linear trend indicates that changes in probability increase with NPI. All the results confirm the findings in the conceptual analysis, that consumers will respond to information cost reduction and benefit by increasing the probability of choosing healthier foods.

We also calculate the annual nutrient intake changes using the simulated changes in probability due to FOP multiplied by market size. We find that FOP decreases per capita calorie intake from RTEC by 0.31%, sugar intake by 2.63%, saturated fat intake by 6.94%, and sodium intake by 1.97%. At the same time, FOP increases fiber intake by 3.24%. This finding supports the argument that using FOP labels decreases the intake of negative nutrients and increases the intake of positive nutrients (Kim, Nayga, and Oral 2000; Bollinger, Leslie, and Sorensen 2011).

Figure 3 graphs the average impact of FOP on the marginal effect of healthfulness by household sizes and frequencies of purchases. A larger impact implies more sensitivity to information cost reduction. The fitted line in the left panel shows a decreasing relationship between the sensitivity to information cost reduction and household sizes. Consumers from smaller households are more sensitive to information cost reduction. The fitted line in the right panel also

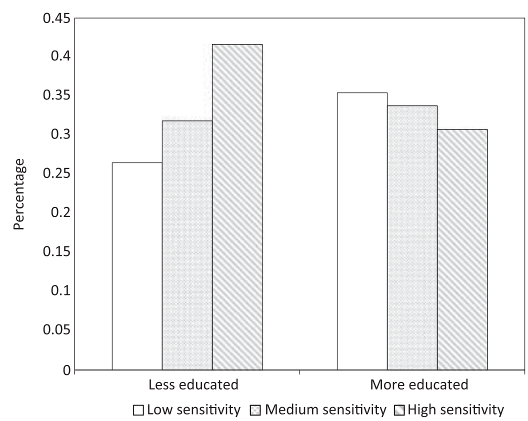

Figure 4. Average impact of FOP on marginal effect of healthfulness by education categories

indicates a decreasing relationship between sensitivity to information cost and frequency of purchases. Consumers who purchase less frequently are more sensitive to information cost reduction. Because the previous literature finds an undetermined relationship between household size and nutritional label usage (Drichoutis, Lazaridis, and Nayga 2005; Govindasamy and Italia 1999; Guthrie et al. 1995; Wang, Fletcher, and Carley 1995), one possible explanation is that smaller households that purchase less frequently pay more attention to the healthfulness attribute of the product when FOP is introduced.

Figure 4 illustrates the impact of FOP on the marginal effect of healthfulness by education levels. We divide the impact into three zones by 1/3 percentiles from the smallest to highest level of sensitivity. More less-educated households are located in the high sensitivity zone compared with the other education categories. For educated households, more are located in the low sensitivity zone. Less-educated households are, on average, 1.3% more responsive to information cost changes than educated households. One possible explanation of this finding is that since education is found to be positively correlated with nutritional information usage (Nayga 1996; Drichoutis, Lazaridis, and Nayga 2006), educated households may already be well-informed about the nutritional information of their choices from the NFP. The addition of FOP does not make significant changes in their choices. The less-educated population benefits more from additional simpler labels.

Conclusions

This study examined the role of nutritional information search cost on the healthfulness of consumer food choices via labeling formats. Using a conceptual analysis and a mixed logit model of consumer choices applied to approximately 2.8 million observations in the RTEC market at the individual level, our main findings are the following.

First, reducing information search cost for nutritional information by introducing simpler FOP labeling formats increases the healthfulness of consumer choices of RTECs. Conceptually, this finding indicates that a lower information search cost increases the reservation healthfulness and ultimately the healthfulness of final product choices. Empirically, we find that the healthfulness attribute weighs 28% more heavily in consumers' choices because of FOP. On average, between 2006 and 2008 for the selected brands, the introduction of front-of-package labeling increased the probability of a consumer choosing a healthy RTEC by 3.49%, and reduced the probability of choosing an unhealthy RTEC by 3.81%. Small but significant decreases in sugar and saturated fat consumption resulted from the introduction of FOP labels.

Second, in terms of consumer heterogeneity, conceptually we find that households with a lower marginal utility of healthfulness *ex ante* are more sensitive to information cost changes. Empirically, we find that households with less education were more responsive to the introduction of FOP labels. Although this may seem counterintuitive, it can be explained by the fact that educated consumers may have higher pre-existing levels of information regarding product healthfulness and may already be making their choices accordingly. Therefore, they are less sensitive to the introduction of simpler labeling formats. We also find that the sensitivity to information cost changes decreases with household size and frequencies of purchases.

In general, the results imply that a private, voluntary label format that provides no additional nutritional information, but which presents the information in a simplified format, may be effective among consumers with less education and smaller families who purchase less frequently. From a policy perspective, this finding demonstrates that the information cost to consumers imposed by the delivery mechanism might be more important than the quantity of information in influencing the healthfulness of consumer food choices.

Some of the limitations of our analysis suggest fruitful avenues for further research. First, in the empirical analysis, we do not consider that changes in information cost may vary across consumers, which is the case when consumers who search for information are "bounded rational" (random errors due to limited attention) or "rationally inattentive" (optimally allocate attention; Manzini and Mariotti 2014; Matejka and McKay 2015). Second, this article relies only on consumer behavior (i.e., the demand side) without considering the response by food companies, which can respond via pricing, promotion, and product strategies, including the introduction of healthier brands that would affect the choice set. Third, we do not consider food-away-from-home in the choice set, where choices are generally less healthy than for food-at-home (Todd et al. 2010). As the 2010 Patient Protection and Affordable Care Act requires all restaurants with 20 or more locations to provide nutritional information on menus, food-away-from-home is bound to weigh more in future debates (Gregory, Rahkovsky, and Anekwe 2014). Finally, whether the results of this study can be extended to other at-home or away-from-home food choices beyond those in our sample is a question that awaits further empirical analysis.

Supplementary Material

Supplementary material is available at http://oxfordjournals.our_journals/ajae/online.

References

Anderson, S.P., and R. Renault. 1999. Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model. *RAND Journal of Economics* 30 (4): 719–35.

Armstrong, M., J. Vickers, and J. Zhou. 2009. Prominence and Consumer Search. *RAND Journal of Economics* 40 (2): 209–33.

Bakos, J.Y. 1997. Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science* 43 (12): 1676–92.

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* 4 (3): 112–27.

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 (1): 1–22.

Berry, S., J. Levinsohn, and A. Pakes. 1995. Automobile Prices in Market Equilibrium. *Econometrica* 63 (4): 841–90.

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 (6): 360–64.

Bollinger, B., P. Leslie, and A. Sorensen. 2011. Calorie Posting in Chain Restaurants. *American Economic Journal: Economic Policy* 3 (1): 91–128.

Castetbon, K., J.L. Harris, and M.B. Schwartz. 2011. Purchases of Ready-to-eat Cereals Vary Across U.S. Household Sociodemographic Categories According to Nutritional Value and Advertising Targets. *Public Health Nutrition* 15 (8): 1456–65.

Chidmi, B., and R.A. Lopez. 2007. Brand-supermarket Demand for Breakfast Cereals and Retail Competition. *American Journal of Agricultural Economics* 89 (2): 324–37.

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 (1): 93–118.

Drichoutis, A.C., P. Lazaridis, and R.M. Nayga. 2006. Consumers' Use of Nutritional Labels: A Review of Research Studies and Issues. *Academy of Marketing Science Review* 9 (9): 1–22.

Govindasamy, R., and J. Italia. 1999. The Influence of Consumer Demographic Characteristics on Nutritional Label Usage. *Journal of Food Products Marketing* 5 (4): 55–68.

Gregory, C., I. Rahkovsky, and T.D. Anekwe. 2014. Consumers' Use of Nutrition Information when Eating Out. Washington DC: U.S. Department of Agriculture, Economic Research Service, Economic Information Bulletin.

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 (5): 385–99.

Guthrie, J.F., J.J. Fox, L.E. Cleveland, and S. Welsh. 1995. Who Uses Nutritional Labeling, and What Effects Does Label Use Have on Diet Quality? *Journal of Nutrition Education* 27 (4): 163–72.

Kiesel, K., J. McCluskey, and S.B. Villas-Boas. 2011. Nutritional Labeling and Consumer Choices. *Annual Review of Resource Economics* 3 (1): 141–58.

Kiesel, K., and S.B. Villas-Boas. 2013. Can Information Costs Affect Consumer Choice? Nutritional Labels in a Supermarket Experiment. *International Journal of Industrial Organization* 31 (2): 153–63.

Kim, S.Y., R.M. Nayga, and C. Oral. 2000. The Effect of Food Label Use on Nutrient Intakes: An Endogenous Switching Regression Analysis. *Journal of Agricultural and Resource Economics* 25 (1): 215–31.

Kiser, E.K. 1998. Heterogeneity in Price Sensitivity and Retail Price Discrimination. *American Journal of Agricultural Economics* 80 (5): 1150–3.

Levy, A.S., and S.B. Fein. 1998. Consumers' Ability to Perform Tasks Using Nutrition Labels. *Journal of Nutrition Education* 30 (4): 210–17.

Liaukonyte, J., N.A. Streletskaya, H.M. Kaiser, and B.J. Rickard. 2013. Consumer

Response to "Contains" and "Free of" Labeling: Evidence from Lab Experiments. *Applied Economics Perspectives and Policy* 35 (3): 476–507.

Manzini, P., and M. Mariotti. 2014. Stochastic Choice and Consideration Sets. *Econometrica* 82 (3): 1153–76.

Matejka, F., and A. McKay. 2015. Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. *American Economic Review* 105 (1): 272–98.

Moorman, C., R. Ferraro, and J. Huber. 2012. Unintended Nutrition Consequences: Firm Responses to the Nutrition Labeling Act. *Marketing Science* 31 (5): 717–37.

Nayga, R.M. 1996. Determinants of Consumers' Use of Nutritional Information on Food Packages. *Journal of Agricultural and Applied Economics* 28 (2): 303–12.

Nevo, A. 2001. Measuring Market Power in the Ready-to-eat Cereal Industry. *Econometrica* 69 (2): 307–42.

Petrin, A., and K. Train. 2010. A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research* 47 (1): 3–13.

Rothman, A.J., R.D. Bartels, J. Wlaschin, and P. Salovey. 2006. The Strategic Use of Gain- and Loss-framed Messages to Promote Healthy Behavior: How Theory Can Inform Practice. *Journal of Communication* 56 (s1): S202–20.

Silberberg, E., and W.C. Suen. 2000. The Structure of Economics: A Mathematical Analysis. New York: McGraw-Hill.

Todd, J.E., and J.N. Variyam. 2008. The Decline in Consumer Use of Food Nutrition Labels, 1995–2006. Washington DC: U.S. Department of Agriculture, Economic Research Service, Technical Report No. 63.

Todd, J.E., L. Mancino, E. Leibtag, and C. Tripodo. 2010. Methodology Behind the Quarterly Food-at-home Price Database. Washington DC: U.S. Department of Agriculture, Economic Research Service, Technical Bulletin No. 1926.

Train, K. 2009. Discrete Choice Methods with Simulation. New York: Cambridge University Press.

Van Wezemael, L., V. Caputo, R.M., Nayga, and G. Chryssochoidis. 2014. European Consumer Preferences for Beef with Nutrition and Health Claims: A Multi-country Investigation Using Discrete Choice Experiments. *Food Policy* 44: 167–76.

Variyam, J.N. 2008. Do Nutrition Labels Improve Dietary Outcomes? *Health Economics* 17 (6): 695–708.

Visschers, V.H., R. Hess, and M. Siegrist. 2010. Health Motivation and Product Design Determine Consumers' Visual Attention to Nutrition Information on Food Products. *Public Health Nutrition* 13 (7): 1099–106.

Volpe, R., A. Okrent, and E. Leibtag. 2013. The Effect of Supercenter-format Stores on the Healthfulness of Consumers' Grocery Purchases. *American Journal of Agricultural Economics* 95 (3): 568–89.

Wang, E., C. Rojas, and C. Bauner. 2015. Evolution of Nutritional Quality in the U.S.: Evidence from the Ready-to-eat Cereal Industry. *Economics Letters* 133: 105–8.

Wang, G., S.M. Fletcher, and D.H. Carley. 1995. Consumer Utilization of Food Labeling as a Source of Nutrition Information. *Journal of Consumer Affairs* 29 (2): 368–80.

Wansink, B., and P. Chandon. 2006. Can "Low-fat" Nutrition Labels Lead to Obesity? *Journal of Marketing Research* 43 (4): 605–17.

Wansink, B., S.T. Sonka, and C.M. Hasler. 2004. Front-label Health Claims: When Less Is More. *Food Policy* 29 (6): 659–67.

Weitzman, M.L. 1979. Optimal Search for the Best Alternative. *Econometrica* 47 (3): 641–54.

Williams, P. 2005. Consumer Understanding and Use of Health Claims for Foods. *Nutrition Reviews* 63 (7): 256–64.

Wolinsky, A. 1986. True Monopolistic Competition as a Result of Imperfect Information. *Quarterly Journal of Economics* 101 (3): 493–512.

Zhen, C., and X. Zheng. 2015. The Effects of Expert Opinion on Consumer Demand for Goods with Credence Attributes: Evidence from a Natural Experiment. Working Paper, Department of Agricultural Economics, North Carolina State University.

Zhu, C., and R. Huang. 2014. Heterogeneity in Consumer Responses to Front-of-package Nutrition Labels: Evidence from a Natural Experiment. Zwick Center for Food and Resource Policy, Working Paper 27, University of Connecticut.