Quantifying the Welfare Effect of Changing Healthy Attributes in the Ready-to-Eat Cereal Market*

Jing Li† and Edward C. Jaenicke ‡

August 9, 2016

Abstract

This paper quantifies the welfare impact of changing healthy attributes of ready-to-eat cereal products. We develop a equilibrium model of production and demand, incorporating the optimal demand choice of consumers facing multiple products in the market and optimal pricing strategy for multi-product firms. We use this model to evaluate how changing healthy attributes in the RTE cereal market affects consumer welfare, producer surplus, and total welfare, taking into account that product prices and consumer demand are endogenous in the equilibrium. We find that improving healthy product attributes has a substantial impact on consumer welfare, producer surplus, and total welfare, but these changes may not necessarily be positive. The results reflect how consumer demand and production costs change when nutritional content changes.

Keywords: *ready-to-eat products; healthy product attributes; consumer welfare; producer surplus*

---

*We thank the Kilts Center for Marketing at Chicago Booth for providing the access to the Nielsen Homescan database.

†Visiting Assistant Professor, School of Business, The University of Hong Kong. Email: `jingli7@hku.hk`.

‡Professor of Agricultural Economics, The Department of Agricultural Economics, Sociology, and Education, The Pennsylvania State University. Email: `tjaenicke@psu.edu`.

1 Introduction

Since 2005, the Dietary Guidelines for Americans recommend that people in the United States should increase their intake of whole grains and dietary fiber, and decrease consumption of refined grains, added sugars, solid fats, and sodium to reduce the risk of heart disease, diabetes, and other chronic health conditions (USDA/HHS, 2005, 2010, 2015). Recent research confirms that U.S. consumers have become increasingly aware of what constitutes a healthy diet (Gregory et al., 2011; Variyam and Smith, 2010; Mancino et al., 2008). To appeal to a more health-conscious consumer base, cereal manufacturers have improved the nutritional profiles of their products. From 2006 to 2012, ready-to-eat (RTE) breakfast cereals manufacturers engaged in product reformulations by reducing sugar and sodium and increasing fiber content. These changes, according to Harris et al. (2012), resulted in an improvement in the nutritional quality of existing cereal brands by 14 percent for childrens cereals, 12 percent for family cereals, and 5 percent for adult cereals, with large RTE breakfast cereal companies contributing the most to this change. Thomas et al. (2013) also confirms this nutritional changes by examining the nutrition data of RTE breakfast cereal products from the two major cereal manufacturers, General Mills and Kellogg, which account for about 62% of the whole RTE cereal market. They find that from 2005 to 2011 the fiber content has increased by about 13.4%, and the sugar and sodium levels decreased by about 7.6% and 11.2% respectively. Whole grain ingredients were found in at least two thirds of the cereals until 2011. Coincidentally RTE cereal sales experienced significant growth around the same time: the two leading firms General Mills and Kellogg experienced a fast growth of sales by about 36% in total, from \$3.9 billion to \$5.3 billion from 2002 to 2010 (General Mills and Kellogg Annual Report to Shareholders, 2003-2010).

In this paper, we evaluate the welfare effect of the aforementioned nutritional improvement of RTE cereals to consumers and manufactures due to the resulting change of consumer demand and associated production costs. More specifically, we quantify the impact of the nutritional improvement on consumer surplus, producer surplus, and total surplus, taking into account that market prices and consumer demand are endogenous in equilibrium. To do so, we extend the demand model of Li et al. (2016) to an equilibrium model by incorporating manufacturers' optimal pricing strategy. The model captures how the nutritional improvement influences the equilibrium prices for each cereal products and accordingly the consumer demand at the new prices.

In order to quantify the effect of the nutritional improvement of the RTE cereals, we first estimate a censored demand system, which is derived from a modified linear approximate almost idea demand system(LA/AIDS) model, using micro-level data following Li et al. (2016). The demand estimation applies the distance metric (DM) approach following Pinkse et al. (2002), Pinkse and Slade (2004), Rojas (2008), and Bonanno (2013), which contributes to a simplified estimation of a demand system and accounting for the impact of product differences in the key attributes. Specifically, estimating a demand system with numerous products would easily encounter large burden of computation due to dimensionality problem. Using micro-level data, which become increasingly popular due to its comprehensive information, can substantially aggravate this problem because micro-level data are usually censored with many zero purchases.[1] The distance metric approach simplifies the censored demand system into a stacked, single equation, which is much easier to implement in estimation when dealing with many-zero-purchase issues in micro-level data such as Nielsen Homescan data set.

We then aggregate the individual consumer demand to market demand, and solve firms' optimal pricing strategy to recover the marginal production costs for each product using information on observed consumer demand, prices, and other observed factors shifting consumer demand and production costs, without actually observing the production data following Nevo (2001). Equipped with the demand and cost functions, we then conduct a series of counterfactual experiments to evaluate how nutritional improvements of RTE cereal products affect consumer surplus, producer surplus, and total surplus, through the channels of resulting equilibrium prices and demand.

The empirical results show that improving health quality of product attributes has a substantial impact on consumer welfare, producer surplus, and total welfare, through their effect on consumer demand, production costs, and the resulting equilibrium prices. But the effects may not necessarily be positive. We find that changing the first ingredient cereal products to whole grain increases consumer welfare, producer surplus, and total welfare. In contrast, improving health quality of product attributes by reducing fat and/or sugar actually reduces them. This is because that consumers actually prefer cereal products with high content of sugar and fat—although they are treated as unhealthy—probably because the taste is better or many consumers do not fully realize the health risk associated with high intake of sugar and fat. The effect of changing fiber content has but a small effect on consumer welfare, producer surplus, and total welfare. These welfare effects reflect the equilibrium results from consumer demand and production costs, associated with the change of nutritional contents.

This paper is related to several papers trying to quantify the impact of product reformulations. For example, Feenstra (1993) find that after quality upgrading for the imported Japanese cars in the United States, the deadweight loss becomes very large. Beatty et al. (2014) find that changes in food formulation account for a substantially large percentage of the dietary improvement among adults in the United States. Griffith et al. (2014) find that consumers' recent decline of intake of dietary salt in UK is entirely attributable to product reformulation by firms. If we focus on the RTE cereal market, we can only observe quite a few. Nevo (2003) conducted a welfare analysis due

[1] For more literature review on solving both dimensionality problem and censoring issue, and a review of distance metric method, please refer to chapter one.

to the quality change and new product introduction in the ready-to-eat breakfast cereal industry during the year from 1988 to 1992. However the results are mostly negative probably due to few products really involve in the reformulation during that time or a very small set of products is considered in his research. In contrast, Mancino et al. (2008) track the trend of consumers purchase of five whole-grain food products from 1998 to 2006, including RTE breakfast cereal. They find that consumers' dietary changes to meet the recommended dietary guideline can attribute to food manufacturers' new product introduction or product reformulation by increasing the content of whole grain in these products led by the 2005 dietary guidelines.

The rest of this paper is organized as follows. Next we will first introduce the data and variable definitions used in the estimation. In section three, we talk about the demand model, and the counterfactual setups by constructing manufacturers' optimization in profit. Section four will show the estimated results, as well as the changes of consumer and manufacturers' surplus. The last section will be a conclusion of this chapter.

2 Data and Variable Definition

2.1 Data

Since the data used in this paper is from the same source as that in chapter one, we only briefly review the data and summarize the main variables used in this this. The first data source is the Nielsen Homescan dataset, which contains detailed information for household level purchases, including the quantity, price, and other product information for each purchase trip. It also contains detailed social demographics for each household which can be matched to each household-level purchase record. This dataset encompasses households living in all 48 contiguous States and the District of Columbia; the households are grouped into 52 metropolitan market areas, plus an additional area that covers the rest of the United States.

In this paper, we focus on the top 20 national cereal brands with the highest number of observed purchase records and sales in this data sets. The total annual sales of these brands account for about 40 percent of the total RTE cereal market in this year. Because a number of cereal brands are targeted to children, we further filter our sample to focus on households with at least one child aged less than 18 years old. Finally, to further reduce the percentage of observed zero purchases, we limit our sample to those household who have at least 15 purchases of RTE cereals over a three-year period. As a result, our data sample extracted from the Nielsen Homescan dataset is constituted by the purchase records from 3,269 households with kids over the top 20 national-brand breakfast cereals in 2004. All the purchase records are yearly aggregated by household and brand.

Next, we match this dataset with brand level product characteristics based on two additional data sources. The first is the USDA National Nutrient Database for Standard Reference. It contains information on over 100 different nutrients and detailed description for each food product. All the nutrient values in this dataset are recorded based on a 100g edible portion. For the purpose of this research, we use several nutrients that are key components for the RTE breakfast cereals, including fiber, sugar, fat, etc. Since cereal manufacturers usually add extra content for some minerals to attract more consumers, we also include the number of vitamins added to fortify the cereal products, which is calculated by using the product description information of the five vitamin components in this dataset.[2]

In addition, since whole grain is commonly accepted as a healthy ingredient in the RTE breakfast cereals, we collected the information of whole-grain type for each cereal product from the official website of each manufacturer. To assist this data collection, we use the "Wayback Machine" website that archives many prominent websites including the main page of RTE cereal manufacturers. Another two attributes, extra flavor added and artificial color added, are also collected by reading the nutrition labels collected in this method. Finally, these three datasets are matched together using each product name and the UPC descriptions.

2.2 Variables for Estimation

Price ($/oz) is calculated by dividing the total dollars paid by the total number of ounces of a given brand bought by a household in each year. However, because the Nielsen Homescan database only reports prices for purchased products, we imputed the missing prices for the non-purchased products by using the average scantrack market price (Arnade et al., 2008). The quantity used in calculating each household's expenditure share is converted to total ounce purchased by each household in a year.

We control for the following demographic variables constructed by using the Nielsen Homescan dataset, including household size (HHSIZE), household income (HHINC), marital status (MARRIED), highest educational attainment (high school, HIGHSC; some college, SOMECOL) of a household head (or the maximum if there is more than one head), highest age (AGEHIGH) of a household head (or the maximum if there is more than one head), and a binary indicator for ethnicity of the household (WHITE, HISPANIC). Summary statistics of these variables are listed in table 1

Following chapter one, we focus on these nutrients: total sugar (SUGAR), total fat (FAT), total dietary fiber (FIBER), and the number of vitamins added to fortify the product (NVF). The content of the first three nutrients is measured as grams per 100 grams of cereal. We construct

[2] The five vitamin components are vitamin A, vitamin D, vitamin E (total ascorbic acid), vitamin B-6, and vitamin B-12.

the variable NVF by summing five binary vitamin-fortification variables: vitamin A, vitamin D, vitamin E (total ascorbic acid), vitamin B-6, and vitamin B-12. For each vitamin considered, a binary variable is given the value one if the brand is enriched with an extra amount of this vitamin according to the reference value from the Recommended Dietary Allowances (RDAs) from the U.S. Food and Drug Administration (USDA/ARS, 2012).

By using the information collected from the Wayback Machine, we are able to create the following binary variables: whole-grain cereals (WG), identifying products whose first ingredient in the ingredients list is whole grain; added flavors (ADFLAVOR), capturing products that report added flavors in the ingredient list; and colors added (ADCOLOR), indicating products that report artificial coloring ingredients in the ingredient list. Also, according to the cereal facts report by Yale University's Rudd Center for Food Policy and Obesity (Harris et al., 2009), we create another binary variable, KIDS, which identifies products targeted to children. In addition, we construct another variable average package size (AVESIZE), by using Homescans information on package size for each product.

All the product attributes mentioned above are considered to play significant impact affecting consumers choices, which will be used in the demand estimation. On the other hand, if we looking at the production process from the view of cereal manufacturers, all the product characteristics will play a role in affecting the outcome of marginal cost for breakfast cereals, including not only the factors that affect consumers purchases decisions. As a result, besides the product attributes mentioned above, we also control several extra cost affecting factors including the nutrition content for sodium(SODIUM), folate(FOLATE), calcium(Calcium), protein(PROTEIN), carbohydrate(CARB). Besides that, we also include the binary variable identifying whether or not the cereal product has fruit added (FRUIT). Finally, in order to approximate the operation cost for each cereal brand, the average package size (AVESIZE) is used to serve a different purpose here as an approximate of stock cost, which could also affect the production cost due to packaging. Besides, we also include the number of unique stores that selling that cereal product (UNISTORE).

Distance Metric Measure

Consumers' purchase decisions for cereal products are not only determined by the characteristics of the product itself, but also by the relative position to other brands. Therefore, consumers' purchase could vary depending on the relative distance among the cereal options. Following Pinkse et al. (2002), Pinkse and Slade (2004) and Rojas (2008), we construct two distance measures in attributes space to quantify the dissimilarity of each unique product relative to the rest products when the consumer is making a choice.

The first one is the continuous distance measure (δ^c_{jk}), which is calculated using continuous product attributes (z^c_j) such as fiber and sugar by following Rojas (2008). It is defined as a function of the inverse Euclidean distance between any pair of products (j and k) in one-dimensional continuous attribute space. In this paper, we focus on three continuous attributes: fiber, sugar and fat. The measure is defined as follows

$$\delta_{jk}^{c} = \frac{1}{1 + 2\sqrt{\left(z_j^c - z_k^c\right)^2}}, \tag{1}$$

where $c = \{1, \ldots, C\}$ represents the index of the three continuous product attributes. z_j^c and z_k^c represent the content of attribute c for product j and k respectively. By construction, this continuous distance measure is positive between 0 and 1. It is smaller when the attribute-space difference between these two products is large.

The second distance measure is (δ_{jk}^{d}), is constructed based on discrete (binary) product attributes. If both products j and k belong to the same product type (e.g., both are whole-grain or both are produced by the same manufacturer), the distance between these two products for that attribute will be one. It equals zero otherwise.

$$\delta_{jk}^{d} = \begin{cases} 1 & if \ \ |z_j^d - z_k^d| = 0, \\ 0 & if \ \ |z_j^d - z_k^d| \neq 0. \end{cases} \tag{2}$$

Similar to the definition for continuous distance measure, $d = \{1, \ldots, D\}$ represents the discrete product attributes, and z_j^d and z_k^d are the attribute type for product j and k accordingly. Smaller value of δ_{jk}^{d}, which means $\delta_{jk}^{d} = 0$, indicates the two products are located far away from each other in this (binary) attribute space. Four binary attributes are considered in constructing the discrete distance measures: whole-grain, added flavor, manufacturer, and whether a product is targeted to children.

Using these distance measures, we further construct the weighted averages of the log form of cross prices in the estimation ($\sum_{j \neq k} \delta_{jk}^{c} log p_{ikt}$ for continuous attributes; $\sum_{j \neq k} \delta_{jk}^{d} log p_{ikt}$ for discrete attributes). The constructed variables are named as DM_FIBER, DM_SUGAR and DM_FAT, using the continuous distance measure for the continuous attribute mentioned above. Using the discrete distance measures, we constructed another four weighted average variables to take into account of whole-grain type (DM_WG), whether a product targets children (DM_KIDS), added flavors (DM_FLAVOR), and manufacturer (DM_MAKER).

3 A Model of Demand and Optimal Pricing

3.1 Consumer Demand

Since the demand estimation in this paper is based on the one in chapter one, we briefly review the demand model in this subsection. Following Lancaster (1979), consumers are assumed to maximize their utility by choosing the level of product attributes within their budget constraint. Under this framework, we follow Rojas and Peterson (2008) to apply the DM method into Deaton and Muellbauer (1980) LA/AIDS model. Consumer $i \in \{1, 2, ..., I\}$ spends a share of her expenditure w_{ijt} on product $j \in \{1, 2, ..., J\}$ at time $t \in \{1, 2, ..., T\}$, which is specified in the same way as chapter one as follows,

$$w_{ijt} = \alpha_{jt} + \sum_{k=1}^{J} b_{ijk} log p_{ikt} + c_{ij} log \frac{x_{it}}{P_{it}^{L}} + e_{ijt}. \tag{3}$$

where $w_{ijt} = q_{ijt}p_{ijt}/x_{it}$ is the expenditure share for consumer i, purchasing product j in market t. p_{ijt} is the unit price for product j and q_{ijt} is the total purchase quantity for consumer i in a year. $x_{it} = \sum_j q_{ijt}p_{ijt}$ is consumer i's total expenditure of all cereal products at market t. The term $logP_{it}^{L}$ is an approximated log-linear analogue of the Laspeyeres Price index (Moschini, 1995), where $logP_{it}^{L} \approx \sum_{j=1}^{J} w_{ij}^{0}$ and w_{ij}^{0} denotes brand j's base share for consumer i with $w_{ij}^{0} = Y^{-1} \sum_{y=1}^{Y} w_{ijy}$ and $y \in (1, ..., Y)$ represents the year. α_{jt}, b_{ijk}, and c_{ij} are demand parameters to be estimated. The unobserved demand shocks e_{ijt} is assumed to be iid over i, j and t.

To further simplify the estimation and overcome the hurdle of dimensionality issue, we follow Pinkse et al. (2002), Pinkse and Slade (2004), Rojas (2008), and Bonanno (2013) to modify equation 3 by defining each cross-price coefficient $b_{ijk}(j \neq k)$ as a linear function of the relative distance in attribute space between product j and k, $b_{ijk} = g(\delta_{jk})$, which reflects the notion that substitutability is impacted by the relative position of the products in characteristics space. The distance measures can be either continuous, δ_{jk}^{c}, which is constructed from continuous attributes (e.g., grams of dietary fiber, fat, or sugar per 100 grams of cereal) or discrete, δ_{jk}^{d} , constructed from discrete characteristics (e.g., whole grain listed as the first ingredient, or a cereal targeted to children). Both of these two distances are constructed following equation 1 and equation 2. Specifically, the cross-price coefficient $b_{ijk}(j \neq k)$ can be defined as:

$$b_{ijk} = \begin{cases} \sum_{c=1}^{C} \lambda_{ij}^{c} \delta_{jk}^{c} & for\ continuous\ distance, \\ \sum_{d=1}^{D} \lambda_{ij}^{d} \delta_{jk}^{d} & for\ continuous\ distance. \end{cases} \tag{4}$$

Here, both λ_{ij}^{c} and λ_{ij}^{d} are assumed to be the same across product equations and consumers, which means $\lambda_{ij}^{c} = \lambda^{c}$ and $\lambda_{ij}^{d} = \lambda^{d}$. Since the constructed distance measures, shown in equation 1 and equation 2, are symmetric, the Slutsky symmetry restrictions are satisfied, which means $b_{ijk} = b_{ikj}$.

To further assist the simplification of the estimated demand model, we define the constant term (α_{jt}), the own-price coefficient (b_{ijj}), and the coefficient of total expenditure (c_{ij}) to be linear functions of product characteristics (z). Besides that, to capture consumers heterogeneous preferences over cereal products, consumers purchase decisions are also allowed to vary depending on their socio-demographic factors, which are modeled as a demand shifter in this paper. As a result, these three terms are defined as: $\alpha_{jt} = \alpha_0 + \sum_{l=1}^{L} a_l z_{jl}^{\alpha} + \sum_{h=1}^{H} a_h h_{hit}$, where h_{hit} is the h-th demographic variable for consumer i at market t; $b_{ijj} = b_0 + \sum_{m=1}^{M} b_m z_{jm}^{\beta}$; and $c_{ij} = c_0 + \sum_{n=1}^{N} c_n z_{jn}^{\gamma}$. All the z variables, z_{jl}^{α}, z_{jm}^{β}, and z_{jn}^{γ}, represent the characteristics for product j. In order to avoid the multicollinearity, each z_{jl}^{α}, z_{jm}^{β}, and z_{jn}^{γ} may be a separate subset of $j^{'}$s characteristics. Up to this point, the estimated demand system are reduced to a single equation based on all the restrictions imposed. Replacing all these modifications of equation 3, our estimated demand model is shown:

$$\begin{aligned} w_{ijt} \quad = \quad & \alpha_0 + \sum_{l=1}^{L} a_l z_{jl}^{\alpha} + \sum_{h=1}^{H} a_h h_{hit} + \left(b_0 + \sum_{m=1}^{M} b_m z_{jm}^{\beta} \right) log p_{ijt} \\ & + \sum_{c=1}^{C} \left(\lambda_c \sum_{j \neq k} \delta_{jk}^{c} log p_{ikt} \right) + \sum_{d=1}^{D} \left(\lambda_d \sum_{j \neq k} \delta_{jk}^{d} log p_{ikt} \right) \\ & + \left(c_0 + \sum_{n=1}^{N} c_n z_{jn}^{\gamma} \right) log \frac{x_{it}}{P_{it}^{L}} + e_{ijt}. \end{aligned} \tag{5}$$

Since each household could only purchase a small subset among all the cereal products, leaving zero-purchases for most of the rest, the expenditure share (w_{ijt}) as defined above are restricted to be non-negative. To take account of this censored nature of our data, the Tobit model (Tobin, 1958) is used in our empirical estimation. Let latent share, w_{ijt}^{*}, represent the unobserved purchase behavior for consumer i choosing product j in market t. The observed share (w_{ijt}) is assumed to be equal to the latent share (w_{ijt}^{*}) whenever the latent share is above zero:

$$w_{ijt}^{*} = \begin{cases} w_{ijt} & if \ w_{ijt}^{*} > 0, \\ 0 & if \ w_{ijt}^{*} \leq 0. \end{cases} \tag{6}$$

The latent share (w^*_{ijt}) is defined by following equation 5:

$$\begin{aligned} w^*_{ijt} \quad = \quad & \alpha_0 + \sum_{l=1}^{L} a_l z^{\alpha}_{jl} + \sum_{h=1}^{H} a_h h_{hit} + \left(b_0 + \sum_{m=1}^{M} b_m z^{\beta}_{jm} \right) log p_{ijt} \\ & + \sum_{c=1}^{C} \left(\lambda_c \sum_{j \neq k} \delta^c_{jk} log p_{ikt} \right) + \sum_{d=1}^{D} \left(\lambda_d \sum_{j \neq k} \delta^d_{jk} log p_{ikt} \right) \\ & + \left(c_0 + \sum_{n=1}^{N} c_n z^{\gamma}_{jn} \right) log \frac{x_{it}}{P^L_{it}} + \epsilon_{ijt}. \end{aligned} \tag{7}$$

where $\epsilon_{ijt} \sim N(0, \sigma^2)$. Under this specification, the zero-purchase outcomes are assumed to represent a corner solution from a utility maximization process.[3] Given the potential endogeneity of prices, the parameters of this model can be consistently estimated using Newey's two-step estimator (Newey, 1987). The final specification given by 6 and 7 allows us to easily estimate demand for a large number of alternatives and account for censoring data fact as well, as it requires the calculation of only a single integral instead of multiple integrals.[4]

3.2 Firms' Optimal Pricing

Following Nevo (2001), there are F cereal manufacturers in the market and each manufacturer produces a subset of differentiated RTE breakfast cereal products ($\mathscr{F}_f$). Each firm chooses the prices for each one of their products to maximize their total profit over all their produced cereals in each market. The profit maximization problem for each manufacturer f at market t is:

$$\pi_{ft} = \max_{p_{jt}, j \in \mathscr{F}_f} \sum_{j \in \mathscr{F}_f} \left(p_{jt} - mc_{jt} \right) q_{jt}\left(p \right) - C_{jt}, \tag{8}$$

where p_{jt} is the price of product j in market t, mc_{jt} is the marginal cost to supply j in market t, and $q_{jt}(p)$ is the quantity sold in market t for product j. Note that we allow the marginal cost (mc_{jt}) to vary in both product and market to capture the idea of potential market-specific costs to supply the same product, such as heterogeneous wage rate and transportation costs in different markets. The total quantity for product j at market t, $q_{jt}(p)$, is recovered by using the estimated expenditure share as defined in the demand model in last section, $w_{ijt}\left(p\right)$. Given that, we first

[3]There are several reasons (both empirical and theoretical) why we observe zero purchases of a specific brand. First, empirical selection of a time period may cause an observation of zero purchases for a specific product. Second, some argue that a zero purchase outcome is the result of a utility-maximization process, and is caused by a relatively high price, unattractive product attributes, or different preferences of consumers. We generally assume that the second reason fits our case.

[4]Since the demand estimation are the same as in chapter one, here for easy reference we recap the estimation results. For more details about specification and estimation, please refer to chapter one.

calculate each consumer i' purchasing quantity for product j in market t:

$$q_{ijt}(p) \quad = \quad \frac{w_{ijt}(p) * x_{it}}{p_{ijt}}. \tag{9}$$

Then the total quantity of product j sold in market t is the sum of individual purchase quantity from consumers who are in market t:

$$q_{jt}(p) \quad = \quad \sum_{i=1}^{I^t} q_{ijt}(p), \tag{10}$$

where I^t represents the total population in market t. Again, following Nevo (2001), we assume that the prices are the outcome of Nash-Bertrand equilibrium, where each firm maximize their total profit over all its products in each market t. Then the set of first order conditions of profit maximization for each product j for manufacturer f at market t becomes:

$$q_{jt}(p) + \sum_{j \in \mathscr{F}_f} (p_{jt} - mc_{jt}) \frac{\partial q_{jt}}{\partial p_{kt}} = 0 \tag{11}$$

Under Nash-Bertrand competition for multiple-product manufacturers, the ownership matrix Ω^*_{jk} can be defined:

$$\Omega^*_{jk} = \begin{cases} 1 & if \;\; k, j \in \mathscr{F}_f, \\ 0 & if \;\; otherwise. \end{cases} \tag{12}$$

Also, define $\Delta_{jt} = \partial q_{jt} / \partial p_{kt}$, and replacing q_{jt} using equation 9 and equation 10, we can get:

$$\Delta_{jk} = \begin{cases} \sum_{i=1}^{I^t} \frac{x_{it}}{p_{jt}^2} \left[\frac{w_{ijt}(p)}{lnp_{jt}} - w_{ijt}(p) \right] & if \;\; k = j, \\ \sum_{i=1}^{I^t} \frac{x_{it}}{p_{kt}^2} \left[\frac{w_{ijt}(p)}{lnp_{kt}} \right] & if \;\; k \neq j. \end{cases} \tag{13}$$

Then define the matrix $\Omega_{jk} = \Omega^*_{jk} * \Delta_{jk}$ and the first order condition from equation 11 can be rewritten in the vector notation as:

$$p - mc = -\Omega^{-1} q(p) \tag{14}$$

where all variables are vectors for all products j. Given the knowledge of prices, ownership matrix, and demand, we can solve out the vector of marginal costs for all products as

$$mc = p + \Omega^{-1} q(p). \tag{15}$$

Marginal Cost Estimation

We first solve for the marginal costs for all products in all markets using equation 15 given the prices, ownership matrix, and demand observed in the data. Because in the data consumer prices varies due to different reasons, such as using coupon or retailer promotion, we approximate the market prices set by the firms as the average prices in each market. Specifically, the market price for product j at market t, p_{jt}, is calculated using the total expenditure for product j at market t divided by total quantity of product j at market t (q_{jt}):

$$p_{jt} = \frac{x_{jt}}{q_{jt}(p)} = \frac{\sum_{i=1}^{I^t} x_{ijt}}{q_{jt}(p)} \tag{16}$$

where the market-level quantity $q_{jt}(p)$ can be calculated using Equation (10), based on the estimated demand function.

We further assume that marginal costs of production is affected by the characteristics of that product (z_j) and related operational cost factors (o_{jt}).

$$mc_{jt} = f(z_j, o_{jt}) \tag{17}$$

In this paper we consider a list of cost-shifting continuous product characteristics including fiber, sugar, fat, sodium, folate, calcium, protein, carbohydrates, and number of vitamin fortifications. We also consider the discrete product characteristics including whether the product is made of whole grain as first ingredient, targeting to children, whether the product has added fruit, and whether or not the product use artificial colors or add extra flavors. Besides including these detailed list of product characteristics, we also consider the average package size as well as the approximate of retailing coverage, e.g. total number of unique stores, to capture the operational cost for each product j at market t. We assume that the marginal cost is a linear function of all these cost-shifting factors mentioned above.

We estimate the marginal cost function (17), using observed data on product characteristics (z_j) and operational cost factors (o_{jt}) and the recovered marginal cost of production from equation (15). The estimation results are reported in table 5. We find that the marginal cost increases if the firm adds fruit, color, flavor or protein to the product. Switch the product to be whole grain, or kids food also increase marginal costs. Similarly, reducing the product contents of sugar, fat, sodium, folate, carbohydrate actually increases production costs, probably because the firm needs to introduce extra producing procedures in order to reduce these contents. However, increase the content of fiber and calcium will help to reduce the cost of production. We also find that increasing package size also help to reduce the marginal cost per ounce. Finally, if a firm increases its retailing coverage, the marginal cost also get reduced.

Then, using the estimated marginal cost function, we are able to predict the new marginal cost given any change of the product characteristics, which will be used in the counterfactual analysis in the next subsection.

3.3 Couterfactual Setup

In this research, we want to evaluate how consumer welfare, producer surplus, and total welfare will be changed if some or all firms change some of their product attributes. For this purpose, we consider a series of counterfactual scenarios in each of which we allow some or all firms change some of important product attributes to the cereal products they produce. Specifically, we consider eight counterfactual scenarios. In the first two, we allow two major cereal manufacturers, Firm A and Firm B[5], to change all their own products to be made using whole grain as primary ingredient, respectively. In the next three counterfactuals, we allow all firms to increase fiber, reduce sugar, or reduce fat by 10% in each of the experiment respectively.[6] The the last three we consider a combination of increasing fiber, and reducing sugar and/or fat, all by 10%. In all these experiments, we assume that the change of product attributes would lead to a increase of product healthy quality of some or all products. A list of the eight counterfactual experiments is reported in Table 6.

In each of the counterfactual, given the proposed change of product attributes we calculate the welfare change using the following procedures:

- Predict the new marginal cost for each product j in each market t using cost equation 17.
- Based on firms' optimal pricing rule in each market, calculate the new prices for each product in each market following equation 15.
- Calculate the new demand (expenditure share) using the new equilibrium price for each individual calculated in last step. Then compute the aggregate demand in each market.
- Given the new prices and new demand, calculate the new consumer surplus for individual

[5]Due to the agreement of using this dataset, we are not able to disclose the name of firm and products. As a result, we use "Firm A" and "Firm B" to represent the two major cereal manufacturers. Since products in this research come from four cereal manufacturers, the other two are name as "Firm Three" and "Firm Four" accordingly.

[6]Although the USDA has confirmed that the amount of fiber in breakfast cereal products from the top two manufacturers increased 32% in the United States from 2005 to 2011, and the amount of sugar and sodium decrease by about 10% and 14% respectively, Williams (2014) found that there is no significant reduce of intake of sodium for consumers who eat breakfast cereals after a systematic review of over two hundreds published scientific literature from all dates until October 2013 in the Scopus and Medline databases. Besides that, Williams (2014) found that consumers who regularly eat breakfast cereal will have higher percentage of intake of fiber and sugar, and lower percentage of intake of fat. As a result, our counterfactual scenarios will focus on these product characteristics: fiber, sugar, fat, and whole grain.

consumer i for each product j following:

$$\begin{aligned} CS_{ijt} &= x_{it}\left(\alpha_0+\sum_{l=1}^{L}a_l z_{jl}^{\alpha}+\sum_{h=1}^{H}a_h h_{hit}\right)(p_{q_{jt}=0}-p_{new}^{jt}) \\ &+\frac{1}{2}\left[b_0+\sum_{m=1}^{M}b_m z_{jm}^{\beta}-\left(c_0+\sum_{n=1}^{N}c_n z_{jn}^{\gamma}\right)w_{ij}^{0}\right]\left[\left(logp_{q_{jt}=0}\right)^2-\left(logp_{new}^{jt}\right)^2\right] \end{aligned} \quad (18)$$

$p_{q_j=0}$ indicates a corner solution for price when the purchase quantity is zero. [7] Then the aggregate consumer surplus is calculated by summing over all the consumers across all the markets, $CS=\sum_j^J\sum_i^I CS_{ijt}$.

- Given the new prices and marginal costs, calculate the new producer surplus for each firm F. Then calculate the new aggregate producer surplus for all firms:

$$PS=\sum_{f}^{F}\sum_{j\in\mathscr{F}_f}\sum_{t}^{T}(p_{new}^{jt}-mc_{new}^{jt})\frac{\sum_i^{I^t}x_{ijt}}{p_{new}^{jt}} \quad (20)$$

where I^t represents all the consumers who are in market t and $\frac{\sum_i^{I^t}x_{ijt}}{p_{new}^{j}}$ is the total quantity sold for product j at market t.

- Compare the consumer surplus, producer surplus, and total surplus (the sum of the consumer surplus and producer surplus) with their counterpart computed in the data.[8]

4 Empirical Results

In this section, we first recap the result of demand estimation from chapter one. Next, we will focus on the results of counterfactual experiments, looking at the change of product prices and welfare after firms change some or all of their product attributes in the eight counterfactual experiments.

[7] In order to calculate the price when $q_j=0$, we solve for the following equation by assuming the price for all the other products unchanged:

$$p_{q_{jt}=0}=exp\left\{\left(\frac{1}{B-Cw_{ij}^{0}}\right)\left[-(A+D)-C\left(logx_{it}-\sum_{j\neq k}w_{ik}^{0}logp_{ikt}\right)\right]\right\} \quad (19)$$

where $A=\alpha_0+\sum_{l=1}^{L}a_l z_{jl}^{\alpha}+\sum_{h=1}^{H}a_h h_{hit}$, $B=b_0+\sum_{m=1}^{M}b_m z_{jm}^{\beta}$, $C=c_0+\sum_{n=1}^{N}c_n z_{jn}^{\gamma}$, and $D=\sum_{c=1}^{C}\left(\lambda_c\sum_{j\neq k}\delta_{jk}^{c}logp_{ikt}\right)+\sum_{d=1}^{D}\left(\lambda_d\sum_{j\neq k}\delta_{jk}^{d}logp_{ikt}\right)$. For simplification of notation, here we remove the sub-index of i for $P_{q_j=0}$. Similar simplification is done for P_new. However, we should clarify that the p_new for each consumer is actually used by the new average market price based on our calculated new equilibrium price.

[8] Following equation 18 and equation 20, we calculate the consumer surplus and producer suplus from the data. Note that, the subscript for price (p_{new}) should change to p_{data} for clear notation.

4.1 Demand Parameters

The estimated demand parameters from the estimated instrumental-variable (IV) Tobit models are shown in table 4.[9]

The own-price coefficient (LNP) is negative and significant (-0.3516) at the 1% level. Estimated coefficients for interactions between price and product attribute show that households in our sample become less price sensitive if cereal products use whole grain as first ingredient (LNP∗WG = 0.0567), have higher levels of fat (LNP∗FAT = 0.0160) and sugar (LNP∗SUGAR = 0.0027), and have more vitamins fortified (LNP∗NVF = 0.0133). While, households become more price sensitive if a cereal product contains artificial color (LNP∗ADCOLOR = -0.0833). For the results for sugar and fat, consumers show a stronger preferences of taste over health, since reducing the content of sugar and fat will lead to a more sensitive to the price, which is consistent to our result of welfare analysis in the following subsection.

The estimated results for the distance metric terms (DM) reflect how consumers substitute between products, which are closer in attribute space, given the changes of price. The positive and significant coefficient for DM_FIBER (0.0011) and DM_WG(0.0123) indicate that, in response to price increases households in our sample are more likely to switch to another cereal product that has similar fiber content or to anther cereal that is also made with whole grain as primary ingredient. On the other hand, for sugar and fat, which are generally advised as unhealthy food elements, the estimated coefficients are negative and significant (DM_SUGAR = -0.0322; DM_FAT = -0.0203). It can be interpreted that households are more likely to switch to cereals that are farther away in the attribute space of sugar or fat when facing a increase of product price. Again, the reason could be that consumers are making tradeoffs between good taste and healthy diet. For example, given a price increase, consumers who are in favor of more healthful cereals with low level of sugar or fat can be induced to switch to cereals that are less healthy but have much better taste. The positive and significant coefficient of DM_KIDS (0.0098) indicates that households with kids tend to switch to other kids' cereals for their children.

The estimated parameter for the expenditure term (LNEP) is positive and significant at the 1% level (0.0903), showing a positive income effect. The result of interaction terms between expenditure and product attributes show that households become more expenditure sensitive if cereal products use whole grain as the first ingredient (LNEP∗WG = 0.0085). Otherwise, household in our sample become less expenditure sensitive for products with more fat (LNEP∗FAT = -0.0063) or fortified with more vitamins (LNEP∗NVF = -0.007). Overall, these results indicate that demand for whole-grain cereals expands the most as the cereal category expands, while demand for vitamin-fortified or higher fat content cereals contracts as the category expands.

[9]For more detail about the selection of model specification and price instrument, please refer chapter one.

Finally, considering the impact of consumers heterogeneous preference, our results show that households that are white in race (WHITE=0.0072) or are larger in size (HHSIZE=0.0051) tend to purchase more cereal. While households with household head who is higher in age (AGEHIGH) tend to purchase less cereal on average (-0.0021). Package size (AVESIZE) also shows a negative impact on cereal purchases (-0.0017). However, neither marriage status, education, household income, nor Hispanic shifters show evidence of affecting household purchasing decisions on breakfast cereal.

4.2 Counterfactual results

Based on the estimated results of consumer demand, we are able to proceed to conduct counterfactual experiments under eight different scenarios, which capture the impact of major nutrients of the RTE breakfast cereals on the welfare for both consumers and producers. In this subsection, we first report the results of new equilibrium price under each counterfactual scenario and then the welfare change for consumer and producer are reported accordingly.

4.2.1 Counterfactual Price

The price changes in the counterfactual experiments are reported in Table 6. Compared with the data, the new prices in the counterfactual experiments are mostly higher except the case when only reducing fiber by 10%. The change of prices (except the fiber only case) ranges from 0.32% to 1.80%. This implies that, in general, improving healthy attributes leads higher prices the firm would like to charge. The price increases because the production costs rise in order to improve healthy attributes, and/or the consumer demand increases in healthy attributes. For example, switching to whole grain on one hand increases consumer demand, and on the other hand it increases production costs. As a result, manufacturers' optimal price rises when they switch to produce whole grain products. When Firm A switch all of its product lines to whole gain, the average market price increases by 0.77%; when Firm B switches all its products to whole grain, the market average price increases by 1.35%. In contrast, when reducing sugar or fat content, the average market price also rises but due to different reasons. When reducing sugar and fat, consumer demand actually decreases as shown in table 4. At the same time, the production cost rises. These two competing forces shape the final price change. The resulting higher price (1.80% for sugar and 1.29% for fat) suggests that the cost changes dominates the demand changes. However, when increasing fiber content, the market average prices actually decreases slightly. The major reason is that the production costs decrease when increasing fiber content. The price change, though, is quite minimal.

4.2.2 Consumer Welfare

We measure consumer welfare as consumer surplus, and use it to evaluate the change of benefit for consumers when some or all healthy attributes are changed by the manufacturers in the RTE cereal market. When some or all manufacturer improves the healthy attributes of their products, as shown in the eight counterfactual experiments, the prices in general increases. At the same time, the demand may also be changed. These two forces together lead to a change of consumer welfare in the counterfactual experiments.

We report the results in table 7. The change of consumer surplus in each of the counterfactual experiments are reported in column 1. Because consumers have heterogeneous demand, some new consumers may start purchase this reformulated product or some existing consumers may stop purchasing this product when prices and healthy attributes are changed. So we also report the number of purchases in each counterfactual experiment in column 2. In column 3, we also report the change of average consumer surplus per actual purchase.

First we observe that changing first ingredient of cereal products to be whole grain has an positive impact on consumer welfare. When Firm A changes the first ingredient of its cereal products to be whole grain, the aggregate consumer welfare increases by 13.22%. This mainly come from the positive effect on consumer demand. Although the prices rises after the change, the positive effect on consumer demand increases aggregate consumer welfare. Similarly, when Firm B changes the first ingredient of its cereal products to be whole grain, the aggregate consumer welfare is improved by about 12.55%.

Second, when reducing sugar or fact, or increasing fiber—which in fact improves healthy quality of the products—the measured consumer welfare actually decreases. For example, when reduces sugar by 10% for all products in the market, aggregate consumer welfare drops by about 20%. This reflects that, although less sugar is good for health, consumers are not aware of it or the reduction of sugar reduces the taste of the product. But from column 2 and 3 in the table, we see that this drop of aggregate consumer welfare is mainly due to the reduced number of consumers actually purchase cereal products. This is likely because consumers are heterogeneous in their preference on the cereal products. Some consumers may prefer sweet or salty cereal so much. When these two nutrient contents are reduced these consumers would simply stop purchasing the products. However, if we calculate the average consumer surplus per actual purchase, as reported in column 3, consumers who continue to purchase the products after reducing fat and sugar actually get a welfare improvement. These consumers may be those who have healthier eating habit and care about their health conditions a lot, as reflected in their demographic variables. Finally, if we look at the changes of two or more health-related nutrients, consumers would benefit more if manufacturers could increase the health quality of cereal products by changing more kinds of nutrients at the same time.

4.2.3 Producer Surplus

The change of production costs and prices due to the change of healthy attributes also affect producer surplus directly. The impact on producer surplus is very important because eventually it is the manufacturers that determine if the health quality of product attributes will be improved. We report the counterfactual results on producer surplus in table 8. We find that the aggregate producer surplus is improved if either Firm A or Firm B changes the first ingredient of their products to be whole grain. When Firm A makes this change, aggregate producer surplus for all firms on average increases by 17.98%. This number is 12.15% when Firm B changes the first ingredient of its products to be whole grain. This result also supports the current business fact that many cereal manufacturers nowadays continue to use or try to change more of their products to use whole grain as the first ingredient. When all firms increases fiber by 10%, the aggregate welfare also increases slightly at 1.82%, mainly due to the reduced prices in the conterfactual.

In contrast, when reducing sugar, fat, or a combination of them with increasing fiber, there is a large drop of producer surplus, ranging from -12.75% to -25.11%. This might be caused by two potential reasons. First, reducing sugar and fat has a negative impact on consumer demand due to the negative change of taste. Second, it also increases production costs according to the estimation of marginal cost function. Both of them lead to the reduction of producer surplus. This explains why in the market, many cereal products still contain high level of sugar and fat, because it is not profitable for firms although it is commonly advised by government agencies or health policy makers that consuming more sugar or fat is not good for health for people in the United States.

Change of product attributes also changes the competition advantage among different cereal manufacturers, thus it may have different impacts on individual firm. In table 9, we further compare the impact of changing healthy attributes on individual firms. We calculate the change of producer surplus for each firm in each counterfactual scenario, compared with the data. It shows very heterogeneous effect for individual firms. First, while switching to whole grain cereal increases aggregate producer surplus, it has very different effect on different firms. For example, if Firm A switches all of its products to whole grain products, it actually loses some producer surplus. This may happen because it rises the production costs for Firm A, and meanwhile Firm A may lose some consumers to other firms because its prices increase due to higher production cost. However, when Firm B switches all its products to whole gain, it has large gains of producer surplus, while all other firms hurt. The different effects on Firm A and Firm B's producer surplus may come from the different product attributes, the similarity among their own products, and the differences between their products to other products available in the market.

When the sugar and/or fat content of all products in the market are reduced, or a combination with rising fiber, the producer surplus for all firms are reduced. This is because all firms might bear similar change of production costs, and that the relative difference of product attributes does not change much from consumer's perspective, while only observing a increase of price. As a result, all firms suffer a loss of producer surplus in these counterfactual experiments, though the magnitude of the changes differ across firms presumably depending on the product attributes in the data.

4.2.4 Total Welfare

The effect on total welfare is also an important indicator for policy makers, for example the United State Department of Agricultural. So we also calculate the total welfare, as the sum of aggregate consumer surplus and producer surplus. The results are reported in table 10.

First, we find that whole grain has a substantial positive impact on total welfare. If Firm A alone changes the first ingredient of all its cereal products to whole grain, the total welfare can be improved by 14.23%, even thought the production costs and prices will be higher. Similarly, if Firm B changes the first ingredient of all its cereal products to whole grain, the total welfare can be improved by 12.47%. This suggests that policy makers should encourage firms to do so.

In contrast, increasing fiber or reducing sugar and/or fat in fact reduces the measured total welfare, although these changes in fact can help improve health. For example, increasing fiber content in all cereal products in the market by 10% reduces total welfare by about 1.03%. Reducing sugar and fat by 10% in the fourth and fifth counterfactuals reduces total welfare substantially by 19.84% and 17.07%, because both firms and consumers lose after these changes. However, the large drop of total welfare may well come from the fact that most consumers either love sweet and salty food so much so that they do not care about health, or they are not aware of the health risk resulting from high intake of sugar and salt. The policy implication is that policy makers should strengthen education about food health, by helping more people realizes the risk of high sugar and salt in the food, and to encourage manufacturers to use more healthy substitutes to replace sugar and salt to product cereal of similar taste.

5 Conclusion

This paper quantifies the welfare effect of changing healthy attributes in the RTE cereal market based on a equilibrium model of demand and production. We find that improving health quality of product attributes has a substantial impact on consumer welfare, producer surplus, and total welfare. But the changes may not necessarily be positive. This reflects the joint results from consumer demand and production costs, associated with the change of nutritional contents. There are two important policy implications from this study. First, policy makers should encourage more cereal manufacturers to use whole grain as the first ingredient of their products. Second, policy makers should also provide more education to consumer about the health risk associated with high intake of sugar and fat, and encourage cereal manufacturers to develop substitutes of sugar and fat.

We admit that the results shown in this paper still have limitations. First, the impact might be overestimated compared to the actions in the real market of the 20 cereals since we assume a change of some nutritional characteristics, for example increase fiber by 10%, for all the 20 cereals. Second, reformulation for cereal industry is a complex process and it might not be the case that manufacturer only change one nutrition at a time. Our results only approximate a few scenarios interested which might not reflect the whole reformulation cases in this industry. However, our research still provide useful insight from these simulated scenarios that consumers do benefit from these reformulation like changing to whole grain, increasing fiber content. With this purpose in mind, our results will also be instructive to the implementation of health related food policies.

References

Arnade, C., M. Gopinath, and D. Pick (2008): "Brand Inertia in U.S. Household Cheese Consumption," *American Journal of Agricultural Economics*.

Beatty, T. K. M., B.-H. Lin, and T. A. Smith (2014): "Is Diet Quality Improving? Distributional Changes in the United States, 1989-2008," *American Journal of Agricultural Economics*, 96, 769–789.

Bonanno, A. (2013): "Functional Foods as Differentiated Products: the Italian Yogurt Market," *European Review of Agricultural Economics*, 40, 45–71.

Deaton, A. and J. Muellbauer (1980): "An Almost Ideal Demand System," *The American Economic Review*, 70, 312–326.

Feenstra, R. C. (1993): "Measuring the Welfare Effect of Quality Change: Theory and Application of Japanese Autos," NBER Working Paper No. 4401. URL: http://www.nber.org/papers/w4401.pdf.

Gregory, C., T. A. Smith, and M. Wendt (2011): "How Americans Rate Their Diet Quality: An Increasingly Realistic Perspective," Economic Information Bulletin No. (EIB-83) 23 pp, URL:http://www.ers.usda.gov/publications/eib-economic-information-bulletin/eib83.aspx.

Griffith, R., M. OConnell, and K. Smith (2014): "The importance of product reformulation versus consumer choice in improving diet quality," Institute for Fiscal Studies, IFS Working Papers: W14/15. URL: http://www.ifs.org.uk/uploads/publications/wps/wp201415.pdf.

Harris, J. L., M. B. Schwartz, and K. D. Brownell (2009): "Evaluating the Nutrition Quality and Marketing of Childrens Cereals," RUL: http://cerealfacts.org/media/Cereal˙FACTS˙Report˙2009.pdf.

Harris, J. L., M. B. Schwartz, K. D. Brownell, V. Sarda, C. Dembek, , C. Munsell, C. Shin, A. Ustjanauskas, and M. Weinberg (2012): "Limited Progress in the Nutrition Quality and Marketing of Children's cereals," URL: http://www.cerealfacts.org/media/Cereal˙FACTS˙Report˙2012˙7.12.pdf.

Lancaster, K. (1979): *Variety, Equity, and Efficiency: Product Variety in An Industrial Society*, New York: Columbia University Press.

Li, J., E. C. Jaenicke, T. D. Anekwe, and A. Bonanno (2016): "Demand for Ready-to-Eat Cereals with Household-level Censored Purchase Data and Nutrition Label Information: A Distance Metric Approach," .

Mancino, L., F. Kuchler, and E. Leibtag (2008): "Getting consumers to eat more wholegrains: The role of policy, information, and food manufacturers," *Food Policy*, 33, 489496.

Moschini, G. (1995): "Units of Measurement and the Stone Index in Demand System Estimation," *American Journal of Agricultural Economics*, 77, 63–68.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307–342.

——— (2003): "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems," *The Review of Economics and Statistics*, 85, 266–275.

Newey, W. K. (1987): "Efficient estimation of limited dependent variable models with endogenous explanatory variables," *Journal of Econometrics*, 36, 231–350.

Pinkse, J. and M. E. Slade (2004): "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review*, 48, 617–643.

Pinkse, J., M. E. Slade, and C. Brett (2002): "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70, 1111–1153.

Rojas, C. (2008): "Price Competition in U.S. Brewing," *The Journal of Industrial Economics*, 56, 1–31.

Rojas, C. and E. B. Peterson (2008): "Demand for Differentiated Products: Price and Advertising Evidence from the US Beer Market," *International Journal of Industrial Organization*, 26, 288–307.

Thomas, R. G., P. R. Pehrsson, J. K. Ahuja, E. Smieja, and K. B. Miller (2013): "Recent trends in ready-to-eat breakfast cereals in the U.S," *Procedia Food Science*, 2, 20–26.

Tobin, J. (1958): "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, 24–36.

USDA/HHS (2005): "Dietary Guidelines for Americans, 2005," URL: http://health.gov/dietaryguidelines/2005.asp.

——— (2010): "The Dietary Guidelines for Americans, 2010," URL: http://health.gov/dietaryguidelines/2010/.

——— (2015): "2015-2020 Dietary Guidelines for Americans," URL: http://health.gov/dietaryguidelines/2015/guidelines/.

Variyam, J. N. and T. A. Smith (2010): "Americans are More Realistic About the Quality of Their Diets," *Amber Waves*, 8, 7.

Williams, P. G. (2014): "The Benets of Breakfast Cereal Consumption: A Systematic Review of the Evidence Base," *American Society for Nutrition*, 5, 636S–673S.

Tables and Figures

Table 1: Summary of purchase data and demographics from Homescan dataset

Variable	Description	Mean	Std.
Purchase			
SH	Individual expenditure share for each brand	0.05	0.12
LNP	Average purchase price (cents), in log form	2.71	0.27
Demographics			
Continuous			
HHSIZE	Number of individuals in the household	4.12	1.12
HHINC	Household annual income	21.17	5.26
AGEHIGH	Max age category of household head(s)	5.60	1.51
Discrete		Frequency of one	
WHITE	Household has children with age below		0.77
MARRIED	Highest edu. attainment for female head: high school		0.85
HISPANIC	Highest edu. attainment for female head: some college		0.10
HIGHSC	Highest edu. attainment for female head: college or more		0.15
SOMECOL	Age of household head is equal to or larger than 65		0.26

Table 2: Summary statistics of health/non-health related product attributes

Variable	Description	Mean	Std.
Continuous			
FAT	Total fat (g/100 g)	2.92	2.39
SUGAR	Total sugars (g/100 g)	27.35	13.30
FIBER	Total dietary fiber (g/100 g)	5.46	3.81
NVF	Number of vitamin fortified	2.85	1.88
AVESIZE	Average package size (oz)	18.54	4.68
MAKER	Manufacturer for the RTE breakfast cereal	2.05	0.80
Discrete		Frequency of one	
WG	Product has whole grain as first ingredient		0.35
KIDS	Target to children		0.50
ADFLAVOR	Other flavor ingredient added		0.45
ADCOLOR	Other artificial color added		0.40

Table 3: Summary statistics of recovered marginal cost and product attributes for cost estimation

Variable	Description	Mean	Std.
Continuous			
FAT	Total fat (g/100 g)	2.92	2.39
SUGAR	Total sugars (g/100 g)	27.35	13.30
FIBER	Total dietary fiber (g/100 g)	5.46	3.81
NVF	Number of vitamin fortified	2.85	1.88
SODIUM	Sodium (mg$*$0.01/100 g)	6.08	1.97
FOLATE	Folate, total (g$*$0.01/100 g)	4.76	3.08
CALCIUM	Calcium (mg$*$0.01/100 g)	1.06	1.43
PROTEIN	Protein (g/100 g)	8.21	4.38
CARB	Carbohydrate, by difference (g/100 g)	82.44	5.64
UNISTORE	Number of unique stores selling product $j*0.01$	2.97	0.42
AVESIZE	Average package size (oz)	18.54	4.68
Discrete		Frequency of one	
WG	Product has whole grain as first ingredient		0.35
KIDS	Target to children		0.50
FRUIT	Add fruit in cereal		0.30
ADFLAVOR	Other flavor ingredient added		0.45
ADCOLOR	Other artificial color added		0.40

Table 4: Parameter estimates of consumer preference

Variables	Coefficient		Standard errors
HHSIZE	0.0051	* * *	0.0015
WHITE	0.0072	* * *	0.0044
MARRIED	-0.0018		0.0050
HISPANIC	-0.0054		0.0060
AGEHIGH	-0.0021	*	0.0011
HIGHSC	0.0035		0.0049
SOMECOL	0.0063		0.0039
HHINC	-0.0002		0.0004
AVESIZE	-0.0017	* * *	0.0002
LNP	-0.3516	* * *	0.0486
LNP∗WG	0.0567	* * *	0.0192
LNP∗FAT	0.0160	* * *	0.0026
LNP∗NVF	0.0133	* * *	0.0051
LNP∗SUGAR	0.0027	* * *	0.0001
LNP∗ADCOLOR	-0.0833	* * *	0.0053
DM_SUGAR	-0.0322	* * *	0.0028
DM_FAT	-0.0203	* * *	0.0012
DM_FIBER	0.0011	**	0.0006
DM_WG	0.0123	* * *	0.0024
DM_KIDS	0.0098	* * *	0.0027
DM_MAKER	-0.0007		0.0032
DM_ADFLAVOR	0.0016		0.0015
LNEP	0.0903	* * *	0.0095
LNEP∗WG	0.0085	*	0.0044
LNEP∗FAT	-0.0063	* * *	0.0009
LNEP∗NVF	-0.0070	* * *	0.0023
FIRM A	-0.0926	**	0.0395
FIRM B	-0.0359		0.0782
FIRM C	-0.0703	* * *	0.0258
CONSTANT	0.2208	**	0.1013
VIF	20.00		
F-test	2854.51		
Endog(p-value)	0.0000		
ALN Test(p-value)	0.2359		

[1] *, **, and * * *represent 10%, 5% and 1% significance levels, respectively.

Table 5: Estimated parameter for marginal cost

Variables	Coefficient		Standard errors
WG	1.0261	* * *	0.3977
FIBER	-0.3405	* * *	0.0580
SUGAR	-0.0634	* * *	0.0132
FAT	-0.1853	**	0.0761
SODIUM	-0.6870	* * *	0.0488
FOLATE	-0.2099	* * *	0.0507
CALCIUM	-0.2733	* * *	0.0881
KIDS	0.5961	**	0.2732
NVF	-0.5510	* * *	0.0813
PROTEIN	0.0344		0.0561
CARB	-0.2469	* * *	0.0558
FRUIT	0.9723	* * *	0.2793
ADCOLOR	0.1323		0.5177
ADFLAVOR	1.2830	* * *	0.2286
UNISTORE	-4.0071	* * *	0.2284
AVESIZE	-0.0545	* * *	0.0037
CONSTANT	65.1889	* * *	5.7250

[1] *, **, and * * *represent 10%, 5% and 1% significance levels, respectively.

Table 6: Average market price under each counterfactual scenarios

Counterfactual Scenarios	Average Price (cent/oz) [1]		
	Before	After	Changes (%)
Convert to whole grain for Firm A products	15.53	15.65	0.77
Convert to whole grain for Firm B products		15.74	1.35
Fiber 10% ↑		15.52	-0.06
Sugar 10% ↓		15.81	1.80
Fat 10% ↓		15.73	1.29
Fiber 10% ↑ & Sugar 10% ↓		15.66	0.84
Fiber 10% ↑ & Fat 10% ↓		15.58	0.32
Fiber 10% ↑ & Sugar 10% ↓ & Fat 10% ↓		15.72	1.22

[1] the first and second row means Firm A or Firm B change all their products to be made with whole grain as first ingredient.
[2] No. of purchase means positive purchase record for product j.
[3] The data used in this paper is by household by product in 2004. That means that each household has only one purchase record for a product j.

Table 7: Consumer welfare under each counterfactual scenario

Counterfactual Scenarios	Consumer Surplus (%)	No. of Purchases Before	No. of Purchases After	Average Surplus per Purchase (%)
Convert to whole grain for Firm A products	13.22	116060	154828	-15.13
Convert to whole grain for Firm B products	12.55		110777	17.92
Fiber 10% ↑	-1.79		119375	-4.52
Sugar 10% ↓	-20.03		47327	96.10
Fat 10% ↓	-17.59		56096	70.49
Fiber 10% ↑ & Sugar 10% ↓	-19.36		48549	92.77
Fiber 10% ↑ & Fat 10% ↓	-16.88		58493	64.92
Fiber 10% ↑ & Sugar 10% ↓ & Fat 10% ↓	-31.08		35863	123.04

Table 8: Percentage change of producer surplus under each counterfactual scenario

Counterfactual Scenarios	Producer Surplus (%) [1]
Convert to whole grain for Firm A products	17.98
Convert to whole grain for Firm B products	12.15
Fiber 10% ↑	1.82
Sugar 10% ↓	-19.11
Fat 10% ↓	-15.11
Fiber 10% ↑ & Sugar 10% ↓	-16.88
Fiber 10% ↑ & Fat 10% ↓	-12.75
Fiber 10% ↑ & Sugar 10% ↓ & Fat 10% ↓	-25.11

Table 9: Producer surplus by firm under each counterfactual scenarios

Counterfactual Scenarios	Firm A	Firm B	Firm C	Firm D
Convert to whole grain for Firm A products	-24.71	82.05	182.86	205.52
Convert to whole grain for Firm B products	-35.37	131.32	-40.79	-31.79
Fiber 10% ↑	0.55	6.93	-12.90	0.96
Sugar 10% ↓	-16.24	-21.51	-45.14	-27.36
Fat 10% ↓	-14.19	-13.17	-39.81	-23.80
Fiber 10% ↑ & Sugar 10% ↓	-14.51	-19.26	-35.67	-23.07
Fiber 10% ↑ & Fat 10% ↓	-12.48	-10.58	-29.65	-19.20
Fiber 10% ↑ & Sugar 10% ↓ & Fat 10% ↓	-22.47	-28.03	-44.18	-32.40

Table 10: Percentage change of total surplus under each counterfactual scenario

Counterfactual Scenarios	Total Surplus (%) [1]
Convert to whole grain for Firm A products	14.23
Convert to whole grain for Firm B products	12.47
Fiber 10% ↑	-1.03
Sugar 10% ↓	-19.84
Fat 10% ↓	-17.07
Fiber 10% ↑ & Sugar 10% ↓	-18.83
Fiber 10% ↑ & Fat 10% ↓	-16.01
Fiber 10% ↑ & Sugar 10% ↓ & Fat 10% ↓	-29.81