DEMAND FOR BREAKFAST CEREALS: WHOLE GRAINS GUIDANCE AND FOOD CHOICE

Ariun Ishdorj and Helen H. Jensen

aishdorj@tamu.edu and hhjensen@iastate.edu

THE ECONOMICS OF FOOD, FOOD CHOICE AND HEALTH
1st joint eaae/aaea seminar

2010

Selected Poster Paper
prepared for presentation at the 1st Joint EAAE/AAEA Seminar

"The Economics of Food, Food Choice and Health"
Freising, Germany, September 15 – 17, 2010

DEMAND FOR BREAKFAST CEREALS: WHOLE GRAINS GUIDANCE AND FOOD CHOICE

Ariun Ishdorj and Helen H. Jensen

Abstract

When using household-level data to examine consumer demand it is common to find that consumers purchase only a subset of the available goods, setting the demand for the remaining goods to zero. Ignoring such censoring of the dependent variables can lead to estimators with poor statistical properties and estimates that lead to poor policy decisions. In this paper we investigate household demand for four types of breakfast cereals, such as whole grain ready-to-eat, non-whole grain ready-to-eat, whole grain hot and non-whole grain hot cereals, using a censored Almost Ideal Demand System (AIDS) and estimate the parameters of the model via Bayesian methods. Using household level scanner data (ACNielsen Homescan) we find that demand for all types of breakfast cereals is inelastic to changes in prices. The expenditure elasticity is slightly above unity for the whole grain ready-to-eat cereals suggesting that as the expenditure on cereals increases households will allocate proportionally more on whole-grain ready-to-eat cereals and less on other cereals.

Keywords: AIDS model, Bayesian econometrics, censored, cereals, whole grains.
JEL Classification: C11; C34; D12.

1 Introduction

The U.S. Department of Health and Human Services (USDHHS) and the U.S. Department of Agriculture (USDA) have published the Dietary Guidelines for Americans since 1980. The Guidelines provide dietary recommendations to aid the development of nutrition programs and to help and encourage consumers to choose diets that meet their nutritional needs and improve their health. The Guidelines are revised every 5 years based on findings from available research. The 2005 Dietary Guidelines for Americans put new emphasis on whole grain consumption by recommending consumption of at least three 1-ounce-equivalent [1] servings of whole grains [2] per day. The proposed 2010 Dietary Guidelines currently under review provide similar guidance by encouraging a diet that emphasizes whole grains, among other foods. In the Guidelines (2005), whole grains are described as follows: "Whole grains, as well as foods made from them, consist of the entire grain seed, usually called the kernel. The kernel is made of three components - the bran, the germ and the endosperm. If the kernel has been cracked, crushed, or flaked, then it must retain nearly the same relative proportion of bran, germ, and endosperm as the original grain in order to be called whole grain" (US DHHS and USDA 2005). Consumption of diets high in whole grains have been reported to have a number of beneficial health effects including reduced risk of cancer (Jacobs, et al. 1998), cardiovascular disease (Truswell, 2002; Liu et al. 1999), diabetes (Fung et al. 2002; Liu et al. 2000), blood pressure (Hallfrisch et al. 2003) and cholesterol (Lumpton, et al. 1994). For more extensive review, see the 2010 Dietary Guidelines for Americans.

The U.S. Food and Drug Administration (FDA), which regulates U.S. nutrition la-

[1] In general, one-ounce slice of bread; one cup of ready-to-eat cereal, or $\frac{1}{2}$ cup of cooked rice, cooked pasta, or cooked cereal can be considered as one-ounce-equivalent from the grains group (http://www.mypyramid.gov).

[2] see Table 1 for the list of whole grains

beling of most foods and authorizes the use of nutrient and health claims, has allowed three health claims related to grain intakes (FDA, 2008). A specific claim for whole grain foods allows the statement that diets rich in whole grain foods and other plant foods and low in total fat, saturated fat, and cholesterol may reduce the risk of heart disease and some cancer. The release of the 2005 Dietary Guidelines and FDA's consideration of health-related claims gave whole grain product manufacturers the opportunity to differentiate their products from refined grain products and the incentive to produce more whole grain products or reformulate the existing products to meet the whole grain requirements. While FDA has no mandatory labeling requirements regarding whole grains, manufacturers can use nutrient labels such as "100 percent whole grain" or "10 grams of whole grain" on the label of their products as long as the statements are not false or misleading (FDA, 2008).

Table 1: Examples of Whole Grains

Brown rice
Buckwheat
Bulgur (cracked wheat)
Millet
Popcorn
Quinoa
Sorghum
Triticale
Whole-grain barley
Whole-grain corn
Whole-oats/oatmeal
Whole rye
Whole wheat
Wild rice

Source: Dietary Guidelines for Americans.

Mandatory labeling provides greater information and therefore more informed con-

sumer choices. However, in the absence of mandatory labeling it is common for third-party labeling service to emerge. In the case of grain products, the Whole Grain Council (WGC), a nonprofit organization, promotes consumption of whole grains through a packaging symbol, a Whole Grain Stamp [3], indicating whole grain content. The Stamp serves as a tool to help consumers easily indentify whole grain products.

Although the lack of clear labeling makes it more difficult for consumers to identify whole grain food products, the availability and consumption of whole grain products are likely to increase (Buzby, Farah and Volke 2005). Policymakers use recommendations from the 2005 Dietary Guidelines in the development of food program guidance. One example is the recently revised food packages for the Supplemental Nutrition Program for Women, Infants and Children (WIC), which include provisions to allow participants to obtain whole grain products effective in 2009.

There are relatively few recent studies of grain consumption. Evidence from food intake surveys indicates that Americans consume less whole grain than recommended. On average, individuals were eating 10 servings of grains a day in 2003, more than recommended daily allowance, of which whole grain accounted for just over 1 serving (Mancino and Buzby 2005). Similar results were found by Lin and Yen (2007). Using data from 1994-96 and 1998 Lin and Yen compared grain consumption of individuals by economic and demographic characteristics and found that individuals consumed more than the recommended daily amount of all grain, while consuming only 34 percent of the amount of whole grain recommended by the 2005 Dietary Guidelines. Analysis of 1999-2000 National Health and Nutrition Examination survey (NHANES) data shows only

[3]Two types of stamps can be awarded, based on the product ingredients and amount of whole grains in the food. Products must contain at least 8 grams of whole grain per labeled serving in order to use the basic Stamp and at least 16 grams of whole grain and where all grains are whole grain to the 100 percent Whole Grain Stamp

15 % of all grains consumed by individuals are whole grain, and most whole grains come from crackers and snacks and from cereals. More specifically, whole grain crackers and snacks account for 5 % of the total grains consumed by individuals, where as ready-to-eat cereals account for 3 % (Mancino and Buzby 2005).

Given the public health interest in increased consumption of whole grains, it is important to have a good understanding of basic demand parameters for grain and cereal products. We consider demand for breakfast cereals, one of the major sources of whole grains in the diet, and estimation based on household level data.

When using household-level data to examine consumer food demand, it is common to find that consumers choose only a subset of the available goods, leaving observed demand for some of the goods to be zero. Ignoring such censoring of the dependent variables can lead to estimators with poor statistical properties and estimates that lead to poor policy decisions. Hence, we carefully address the issue of censoring in a demand system framework. There exist a number of estimation procedures that handle this censoring problem (Wales and Woodland 1983; Lee and Pitt 1986). Although theoretically consistent, these approaches suffer from the drawback that in the case of many non-consumed goods for some households, evaluation of multiple integrals is necessary. An alternative approach is an Amemiya-Tobin approach, which is the generalization of Tobin's (1958) limited dependent variable model proposed by Amemiya (1974) and implemented by Wales and Woodland (1983). However, the use of Amemiya-Tobin type estimators is also complicated by the need for evaluating multiple integrals in cases where censoring is severe. Due to the complexity of estimating the models above, a two-step procedure based on the Amemiya-Tobin approach has sometimes been used to estimate censored demand systems (Shonkwiler and Yen (1999)). This method has been widely used in the

applied literature. Although the two-step procedure holds an advantage in its ability to estimate large systems, the two-step procedures are known to be inefficient and overlook the adding-up condition of the observed shares.

A number of papers have used variations of the Amemiya-Tobin approach to deal with the issues of censoring in food demand (e.g. Yen and Roe (1989), Perali and Chavas (2000), Golan, Perloff and Shen (2001), Yen, Kan and Su (2002) and Yen (2005)). Advances in simulation methods that allow approximations of high-dimensional integrals have been used in the estimation of the censored demand system (Yen, Lin and Smallwood (2003), Dong, Gould and Kaiser (2004)).

In this paper we propose a Bayesian procedure for estimating the censored demand system using the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980). Estimating a censored AIDS model with a Bayesian approach avoids the need to evaluate the multiple probability integrals. The marginal distribution of model parameters and latent shares are simulated by numerical methods. Specifically, we fit the model using the Gibbs sampler. Implementation of the Gibbs sampler involves deriving and then iteratively simulating from the conditional posterior distribution of the model parameters. The method developed is used to examine the demand for different types of breakfast cereals. We use data from 2006 ACNielsen Homescan household level scanner data files.

The estimation focuses on cereal (whole grain and other, ready-to-eat and hot) products which form a product group widely consumed in the United States. Lin and Yen (2007) found that breakfast was a good source of whole grain. Individuals consumed 40 percent of whole grain at breakfast, compared with 23 percent at lunch and 17 percent at dinner and the rest provided by snack foods. Although scanner data provide informa-

tion on foods purchased only for at home consumption, cereals are generally purchased in retail food stores, and in case of the breakfast cereals, generally consumed at home. Hence, scanner data are well suited for estimating demand relationships for this product group.

In estimating the demand system for cereal products we assume that demand for cereal is separable from the demand for other goods in the consumer budget. In a multistage budgeting framework, it is usually assumed that consumers first allocate their expenditures to broad aggregate commodity groups. Subsequently, consumer's decisions are based on group expenditures and commodity prices within each group. Hence, by weak separability we focus on a demand structure in which cereal expenditures are allocated to various types of cereals.

The paper is organized as follows. The next section describes the AIDS model and the associated Bayesian posterior simulator. Then data used in the analysis are described, followed by a description of empirical results. The paper concludes with a summary of the findings and the directions for the future research.

2 AIDS Model and Posterior Simulator

2.1 The Model

The AIDS model of Deaton and Muellbauer (1980) can be expressed in the latent expenditure share form as: [4]

$$s^*_{ih} = \alpha_i + \boldsymbol{z_{ih}\delta_i} + \sum_{j=1}^{n} \gamma_{ij} ln\left(p_{jh}\right) + \beta_i ln\left(y_h/P_h\right) + \epsilon_{ih}, \quad i,j = 1,...,n, \quad h = 1,..,H \quad (1)$$

[4]Matrices and vectors are denoted by bold letters.

and

$$s_{ih} = \begin{cases} s^*_{ih} & if \quad s^*_{ih} > 0 \\ 0 & if \quad s^*_{ih} \leq 0 \end{cases} \tag{2}$$

where, s^*_{ih} and s_{ih} are the latent and observed expenditure shares, respectively, for good i of household h, p_{jh} is the price of the jth good, $\mathbf{z_{ih}}$ is a set of household specific characteristics, y_h represents total expenditure of household h on all n goods and P_h is a price index defined as:

$$lnP_h = \alpha_0 + \sum_{i=1}^{n} \alpha_i ln(p_{ih}) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} ln(p_{ih}) ln(p_{jh}). \tag{3}$$

The theoretical properties of the demand function given by equation (1) can be imposed by the following equality restrictions on the parameters [5]

adding-up: $\sum_i \alpha_i = 1, \quad \sum_i \gamma_{ij} = \sum_i \beta_i = \sum_i \boldsymbol{\delta_i} = 0;$

homogeneity: $\sum_j \gamma_{ij} = 0$ and

symmetry: $\gamma_{ij} = \gamma_{ji}, \quad i \neq j, \quad i, j = 1, ..., n.$

For each household h stacking (1) over $i = 1, ...n$ we obtain:

$$\mathbf{s^*_h} = \boldsymbol{\alpha} + \mathbf{Z_h}\boldsymbol{\delta} + \mathbf{ln(p_h)}\boldsymbol{\gamma} + \boldsymbol{\beta}\mathbf{ln(y_h/P_h)} + \boldsymbol{\epsilon_h}, \tag{4}$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \boldsymbol{\delta} = \begin{bmatrix} \boldsymbol{\delta_1} \\ \boldsymbol{\delta_2} \\ \vdots \\ \boldsymbol{\delta_n} \end{bmatrix}, \quad \boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\gamma_1} \\ \boldsymbol{\gamma_2} \\ \vdots \\ \boldsymbol{\gamma_n} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix},$$

[5] Here, we are not imposing the adding up to unity restriction, $\sum_i s^*_{ih} = 1$, on the latent shares.

$$\mathbf{s_h^*} = \begin{bmatrix} s_{1h}^* \\ s_{2h}^* \\ \vdots \\ s_{nh}^* \end{bmatrix}, \quad \mathbf{Z_h} = \begin{bmatrix} \boldsymbol{z_{1h}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{z_{2h}} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{z_{nh}} \end{bmatrix},$$

$$\mathbf{ln(p_h)} = \begin{bmatrix} \boldsymbol{ln(p_{1h})} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{ln(p_{2h})} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{ln(p_{nh})} \end{bmatrix},$$

and

$$\mathbf{ln(y_h/P_h)} = ln(y_h/P_h)\mathbf{i_h},$$

We can rewrite (4) as:

$$\mathbf{s_h^*} = \mathbf{X_h}\boldsymbol{\theta} + \boldsymbol{\epsilon_h}, \tag{5}$$

where $\mathbf{X_h} = [\mathbf{I}\ \mathbf{Z_h}\ \mathbf{ln(p_h)}\ \mathbf{ln(y_h/P_h)}]$ is the $n \times k$ matrix of stacked covariate data, $k = \sum_{i=1}^{n} k_i$, k_i denotes the number of explanatory variables, $\boldsymbol{\theta} = [\boldsymbol{\alpha'}\ \boldsymbol{\delta'}\ \boldsymbol{\gamma'}\ \boldsymbol{\beta'}]'$ is $k \times 1$ vector and $\boldsymbol{\epsilon_h} \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma})$ where Σ is $n \times n$.

Stacking (5) over $h = 1, .., H$ we obtain:

$$\mathbf{s}^* = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{6}$$

where $\mathbf{X}$ is $nH \times k$, $\boldsymbol{\epsilon} \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Omega})$ and $\boldsymbol{\Omega}$ is $\mathbf{I_H} \otimes \boldsymbol{\Sigma}$ matrix.

The AIDS model specified above is a Seemingly Unrelated Regression (SUR) model proposed by Zellner (1962) on the latent data $\mathbf{s}^*$, with the same regressors in each equation. Since the expenditure shares are censored we follow Huang et al.(1987) and

estimate a SUR Tobit model.

To impose the parameter restrictions in the estimation of (6) we follow the method specified in Griffiths, O'Donnell and Tan Cruz (2000). Let J, where $J < k$, be the number of equality restrictions imposed on the parameters of the model, then

$$\mathbf{R}\boldsymbol{\theta} = \mathbf{r}, \tag{7}$$

where $\mathbf{R}$ is $J \times k$ and $\mathbf{r}$ is $J \times 1$.

As an example, suppose we want to estimate the following two equation system:

$$s_1^* = \alpha_1 + \gamma_{11}x_{11} + \gamma_{12}x_{12} + \epsilon_1$$

$$s_2^* = \alpha_2 + \gamma_{21}x_{21} + \gamma_{22}x_{22} + \epsilon_2$$

and the linear restrictions that we want to impose are

$$\sum_i \alpha_i = 1, \quad \sum_i \gamma_{ij} = 0 \quad \text{and} \quad \gamma_{12} = \gamma_{21}.$$

Then $\mathbf{R}\boldsymbol{\theta} = \mathbf{r}$ in this case will be:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \gamma_{11} \\ \gamma_{12} \\ \alpha_2 \\ \gamma_{21} \\ \gamma_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

These restrictions imply that some of the parameters of the model are redundant and can be recovered from the estimated parameters and imposed parameter restrictions. We will rearrange the elements of $\boldsymbol{\theta}$ and partition it into vectors of redundant and free parameters, denoted $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$, respectively, where $\boldsymbol{\theta_1}$ is $J \times 1$ and $\boldsymbol{\theta_2}$ is $(k-J) \times 1$. Accordingly, we partition $\mathbf{X}$ by reordering its columns so that equations (6) and (7) can be written as:

$$\mathbf{s}^* = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon} = \begin{bmatrix} \mathbf{X_1} & \mathbf{X_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta_1} \\ \boldsymbol{\theta_2} \end{bmatrix} + \boldsymbol{\epsilon}, \tag{8}$$

and

$$\mathbf{R}\boldsymbol{\theta} = \begin{bmatrix} \mathbf{R_1} & \mathbf{R_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta_1} \\ \boldsymbol{\theta_2} \end{bmatrix} = \mathbf{r}, \tag{9}$$

where $\mathbf{X_1}$ and $\mathbf{X_2}$ are $nH \times J$ and $nH \times (k-J)$ submatrices of $\mathbf{X}$, respectively, $\mathbf{R_1}$ is $J \times J$, $\mathbf{R_2}$ is $J \times (k-J)$ and $\text{rank}(\mathbf{R_1}) = J$. In this notation the covariate matrix is no longer block-diagonal. For the example mentioned above

$$\boldsymbol{\theta_1} = \begin{bmatrix} \gamma_{12} \\ \alpha_2 \\ \gamma_{21} \\ \gamma_{22} \end{bmatrix}, \quad \boldsymbol{\theta_2} = \begin{bmatrix} \alpha_1 \\ \gamma_{11} \end{bmatrix}, \quad \mathbf{R_1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix}, \quad \mathbf{R_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{X_1} = \begin{bmatrix} x_{12} & 0 & 0 & 0 \\ 0 & 1 & x_{21} & x_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{X_2} = \begin{bmatrix} 1 & x_{11} \\ 0 & 0 \end{bmatrix}.$$

As mentioned earlier, we only need to estimate $\boldsymbol{\theta_2}$, since $\boldsymbol{\theta_1}$ is redundant and can be

recovered from $\boldsymbol{\theta_2}$ and imposed restrictions. Solving for $\boldsymbol{\theta_1}$ from (9) we get:

$$\boldsymbol{\theta_1} = \mathbf{R_1^{-1}}(\mathbf{r} - \mathbf{R_2}\boldsymbol{\theta_2}). \tag{10}$$

By substituting $\boldsymbol{\theta_1}$ into (8) and rearranging terms we get

$$\tilde{\mathbf{s}}^* = \tilde{\mathbf{X}}\boldsymbol{\theta_2} + \boldsymbol{\epsilon} \tag{11}$$

where $\tilde{\mathbf{s}}^* = \mathbf{s}^* - \mathbf{X_1}\mathbf{R_1^{-1}}\mathbf{r}$ and $\tilde{\mathbf{X}} = \mathbf{X_2} - \mathbf{X_1}\mathbf{R_1^{-1}}\mathbf{R_2}$. Thus, (11) is a latent variable SUR model with no restrictions on $\boldsymbol{\theta_2}$.

2.2 The Augmented Posterior

For computational simplicity, we follow Albert and Chib (1993) and treat the latent data $\tilde{\mathbf{s}}^*$ as additional parameters of the model. The augmented posterior $p(\tilde{\mathbf{s}}^*, \boldsymbol{\theta_2}, \boldsymbol{\Sigma}|\mathbf{s})$ is then proportional to

$$\begin{aligned} p(\tilde{\mathbf{s}}^*, \boldsymbol{\theta_2}, \boldsymbol{\Sigma}|\mathbf{s}) &\propto p(\mathbf{s}|\tilde{\mathbf{s}}^*, \boldsymbol{\theta_2}, \boldsymbol{\Sigma})p(\tilde{\mathbf{s}}^*|\boldsymbol{\theta_2}, \boldsymbol{\Sigma})p(\boldsymbol{\theta_2}, \boldsymbol{\Sigma}) && (12)\\ &\propto p(\boldsymbol{\theta_2}, \boldsymbol{\Sigma})\left(\prod_{h=1}^{H} p(\mathbf{s_h}|\tilde{\mathbf{s}}_\mathbf{h}^*)p(\tilde{\mathbf{s}}_\mathbf{h}^*|\boldsymbol{\theta_2}, \boldsymbol{\Sigma})\right) && (13)\\ &\propto p(\boldsymbol{\theta_2}, \boldsymbol{\Sigma})\left[\prod_{h=1}^{H} p(\tilde{\mathbf{s}}_\mathbf{h}^*|\boldsymbol{\theta_2}, \boldsymbol{\Sigma})\left(\prod_{i=1}^{n} p(s_{ih}|\tilde{s}_{ih}^*)\right)\right], && (14)\end{aligned}$$

where

$$p(s_{ih}|\tilde{s}_{ih}^*) = I(s_{ih} = \tilde{s}_{ih}^*)I(\tilde{s}_{ih}^* > c_h) + I(s_{ih} = 0)I(\tilde{s}_{ih}^* \leq c_h),$$

and c_h is the hth element of $-\mathbf{X_1}\mathbf{R_1^{-1}}\mathbf{r}$.

From (4), the sampling density of the latent data, $\tilde{\mathbf{s}}^*$, is given as:

$$p(\tilde{\mathbf{s}}^*_{\mathbf{h}}|\boldsymbol{\theta_2}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{H}{2}} \exp\left(-\frac{1}{2}\sum_{h=1}^{H}(\tilde{\mathbf{s}}^*_{\mathbf{h}} - \tilde{\mathbf{X}}_{\mathbf{h}}\boldsymbol{\theta_2})'\boldsymbol{\Sigma^{-1}}\sum_{h=1}^{H}(\tilde{\mathbf{s}}^*_{\mathbf{h}} - \tilde{\mathbf{X}}_{\mathbf{h}}\boldsymbol{\theta_2})\right) \tag{15}$$

To implement a Bayesian analysis, we must introduce the priors. We assume that the priors are independent and of the conditionally conjugate forms:

$$\boldsymbol{\theta_2} \sim N(\mu_{\boldsymbol{\theta_2}}, \boldsymbol{V}_{\boldsymbol{\theta_2}}) \tag{16}$$

$$\boldsymbol{\Sigma}^{-1} \sim W(\underline{\mathbf{A}}, \underline{\nu}), \tag{17}$$

where W denotes a Wishart distribution (Koop, Poirier and Tobias, 2007, pg. 339).

2.3 The Posterior Simulator

In this section we introduce our posterior simulator for fitting the demand model given by (11) together with the priors in (16)-(17). We use the Gibbs sampling algorithm to iteratively draw values from the posterior distribution of each parameter conditional on other parameters of the model. Those posterior conditionals are enumerated below.

Step 1: $\boldsymbol{\theta_2}|\mathbf{s}, \boldsymbol{\Sigma}$

$$\boldsymbol{\theta_2}|\mathbf{s}, \boldsymbol{\Sigma} \sim N(\mathbf{D}_{\boldsymbol{\theta_2}}\mathbf{d}_{\boldsymbol{\theta_2}}, \mathbf{D}_{\boldsymbol{\theta_2}}), \tag{18}$$

where

$$\mathbf{D}_{\boldsymbol{\theta_2}} = \left(\tilde{\mathbf{X}}'(\boldsymbol{\Sigma^{-1}} \otimes \mathbf{I_H})\tilde{\mathbf{X}} + V_{\boldsymbol{\theta_2}}^{-1}\right)^{-1}$$

$$\mathbf{d}_{\boldsymbol{\theta_2}} = \left(\tilde{\mathbf{X}}'(\boldsymbol{\Sigma^{-1}} \otimes \mathbf{I_H})\tilde{\mathbf{s}^*} + \mathbf{V}_{\boldsymbol{\theta_2}}^{-1}\boldsymbol{\mu}_{\boldsymbol{\theta_2}}\right)$$

Step 2: $\boldsymbol{\Sigma^{-1}}|\boldsymbol{\theta_2}, \mathbf{s}$

$$\boldsymbol{\Sigma^{-1}}|\boldsymbol{\theta_2}, \mathbf{s} \sim W\left(\overline{\mathbf{A}}, \overline{\nu}\right) \tag{19}$$

where

$$\overline{\nu} = H + \underline{\nu}$$

and

$$\overline{\mathbf{A}} = \left[\underline{\mathbf{A}}^{-1} + \sum_{h=1}^{H}\left(\tilde{\mathbf{s}}_{\mathbf{h}}^{*} - \tilde{\mathbf{X}}_{\mathbf{h}}\boldsymbol{\theta_2}\right)\left(\tilde{\mathbf{s}}_{\mathbf{h}}^{*} - \tilde{\mathbf{X}}_{\mathbf{h}}\boldsymbol{\theta_2}\right)'\right]^{-1}$$

Step 3: $\tilde{s}_{ih}^{*}|\mathbf{s}, \boldsymbol{\theta_2}, \boldsymbol{\Sigma}$

From (14) the posterior conditional of $\tilde{\mathbf{s}}_{\mathbf{h}}^{*}$ is multivariate truncated normal. We therefore follow Geweke (1991) and draw each latent, $\tilde{s}_{ih}^{*}$ from a univariate truncated normal density.

Let ω_{ij} denote the (i, j) element of Σ^{-1} and c_h be the hth element of $-\mathbf{X_1}\mathbf{R_1^{-1}}\mathbf{r}$ as defined before. For each household h we can idependently sample each of the n goods, $i = 1, ..., n$ as follows [6]:

$$\tilde{s}_{ih}^{*}|\mathbf{s}, \boldsymbol{\theta_2}, \boldsymbol{\Sigma} \sim TN_{(-\infty, c_h)}\left(\mu_{i|-i}, \omega_{ii}^{-1}\right), \quad \text{if } \tilde{s}_{ih} = 0, \tag{20}$$

where

$$\mu_{i|-i} = \mu_i - \omega_{ii}^{-1}\sum_{i \neq j}\omega_{ji}(\tilde{\mathbf{s}}_{-i}^{*} - \boldsymbol{\mu_{-i}})$$

then repeat for $h = 1, 2, \ldots, H$.

In the above, $TN_{(a,b)}(\mu, \sigma^2)$ denotes a normal density with mean μ and variance σ^2 truncated to the interval (a, b), μ_i is the ith row element of μ, μ_{-i} denotes all the elements

[6] The way the dependent variables are specified in our model it is possible that the observed shares are clustered both at zero and at one. Accounting for the two-sided censoring in the specification of the model is appropriate. However, only 5%, 4%, 3% and 2% of observed shares in our data are clustered at one. Hence, in this analysis we consider only the case when the observed shares are clusterd at zero.

of μ other than μ_i.

The posterior simulator involves iteratively drawing from (18)-(20).

3 A Generated Data Experiment

In this section we conduct a generated data experiment to demonstrate the performance of our posterior simulator. A sample of 10,000 households is generated from the following demand model:

$$
\begin{aligned}
s^*_{1h} &= \alpha_1 + \gamma_{11} ln(p_{1h}) + \gamma_{12} ln(p_{2h}) + \gamma_{13} ln(p_{3h}) + \gamma_{14} ln(p_{4h}) + \beta_1 ln(y_h/P_h) + \epsilon_{1h} \\
s^*_{2h} &= \alpha_2 + \gamma_{21} ln(p_{1h}) + \gamma_{22} ln(p_{2h}) + \gamma_{13} ln(p_{3h}) + \gamma_{24} ln(p_{4h}) + \beta_2 ln(y_h/P_h) + \epsilon_{2h} \\
s^*_{3h} &= \alpha_3 + \gamma_{31} ln(p_{1h}) + \gamma_{32} ln(p_{2h}) + \gamma_{33} ln(p_{3h}) + \gamma_{34} ln(p_{4h}) + \beta_3 ln(y_h/P_h) + \epsilon_{3h} \\
s^*_{4h} &= \alpha_4 + \gamma_{41} ln(p_{1h}) + \gamma_{42} ln(p_{2h}) + \gamma_{43} ln(p_{3h}) + \gamma_{44} ln(p_{4h}) + \beta_4 ln(y_h/P_h) + \epsilon_{4h}
\end{aligned}
$$

where $ln(p_{ih})$ and $ln(y_h/P_h)$ are drawn independently from a $N(0,1)$ and the error terms $[\epsilon_{1h} \quad \epsilon_{2h} \quad \epsilon_{3h} \quad \epsilon_{4h}]'$ are drawn jointly from the multivariate Normal distribution:

$$
\begin{bmatrix} \epsilon_{1h} \\ \epsilon_{2h} \\ \epsilon_{3h} \\ \epsilon_{4h} \end{bmatrix} \overset{iid}{\sim} N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} .5 & -.45\sqrt{.5}\sqrt{.3} & .5\sqrt{.5}\sqrt{.1} & -.35\sqrt{.5}\sqrt{.6} \\ -.45\sqrt{.5}\sqrt{.3} & .3 & -.2\sqrt{.3}\sqrt{.1} & .4\sqrt{.3}\sqrt{.6} \\ .5\sqrt{.5}\sqrt{.3} & -.2\sqrt{.3}\sqrt{.1} & .1 & -.5\sqrt{.1}\sqrt{.6} \\ -.35\sqrt{.5}\sqrt{.6} & .4\sqrt{.3}\sqrt{.6} & -.5\sqrt{.1}\sqrt{.6} & .6 \end{pmatrix} \right]
$$

Some of the variables in our actual data have high degree of censoring. To imitate the actual data as close as possible we generate the data with 30 %, 21 %, 56 % and 70 % of censoring. We fit our model using the posterior simulator described in previous

section, ran the algorithm 100,000 iterations, and discarded the first 30,000 draws as the burn-in period.

Table 2 and Figures 1 and 2 summarize the results of the generated data experiment. We plot the lagged autocorrelations up to order 8 for several selected parameters: γ_{14}, α_2, γ_{33}, γ_{41}, ρ_{12}, ρ_{24}, σ_1^2 and σ_4^2. From the plots we can see that the Gibbs sampler displays good mixing of the parameters.

In Table 2 we report the estimates of the posterior means, standard deviations and probabilities of being positive from the generated data along with their true values. As we can see from the table, all the parameters have been estimated with reasonable accuracy and the estimated results are quite close to their true values.

4 The Data

4.1 Household Data

We use data from the ACNielsen 2006 Homescan survey of households. The data come from a nationally representative sample of U.S. households that scan their purchased foods at home after each shopping occasion using a scanning device and report the results to the collection firm once a week. The dataset includes product modules of dairy department purchase data, dry grocery department purchase data, produce, meat and frozen departments purchase data and a module for random-weight purchase data for the year of 2006. Each product module and the random-weight data includes product codes that identify brand, size, flavor, form, formula, container, style, type and variety. Each food item was represented by a unique UPC or product number. The data also contain information on purchase date, quantity purchased, total expenditures on the

Figure 1: Lagged Autocorrelations for γ_{14}, α_2, γ_{33} and γ_{41}

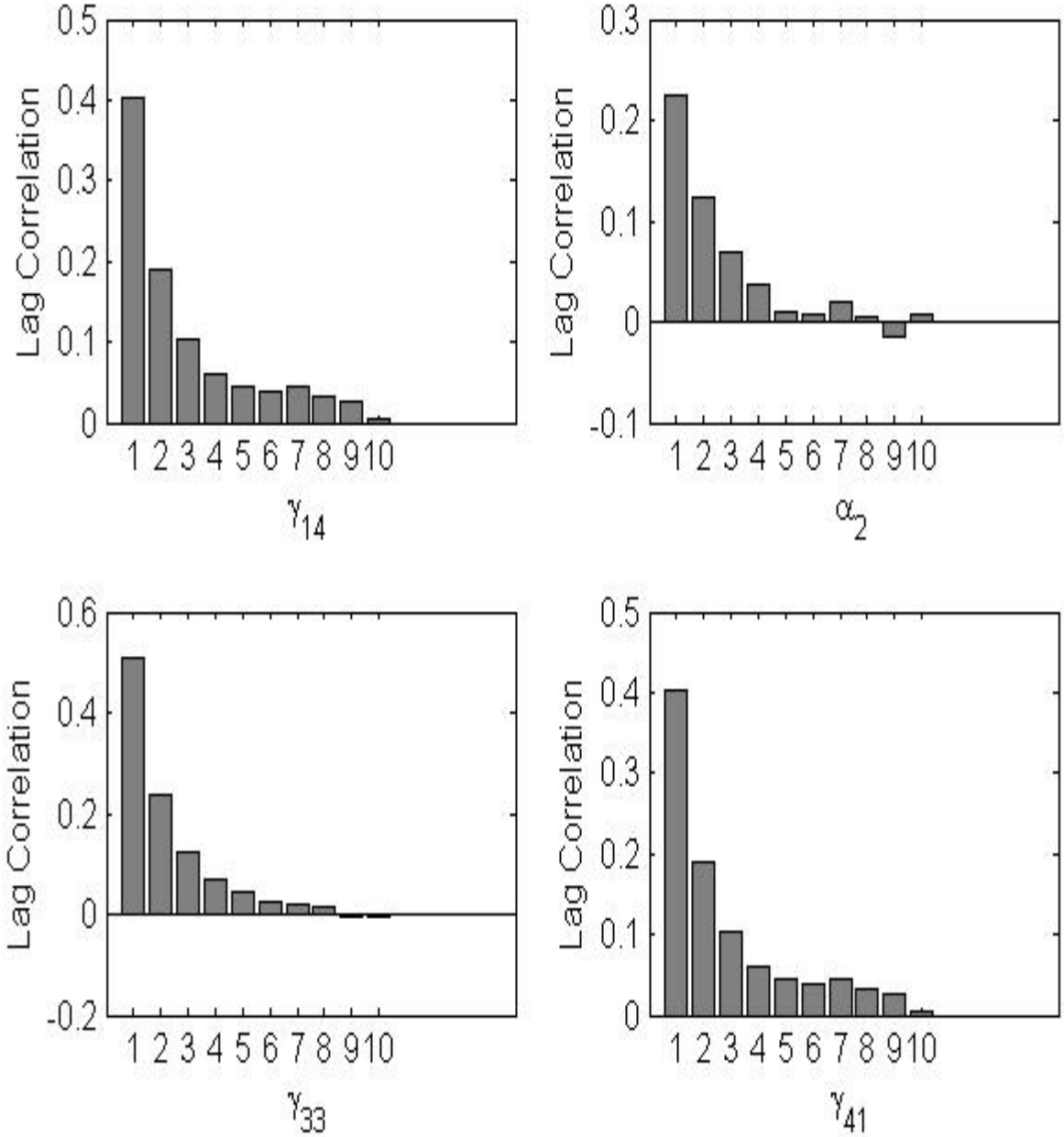

Figure 2: Lagged Autocorrelations for ρ_{12}, ρ_{24}, σ_1^2 and σ_4^2

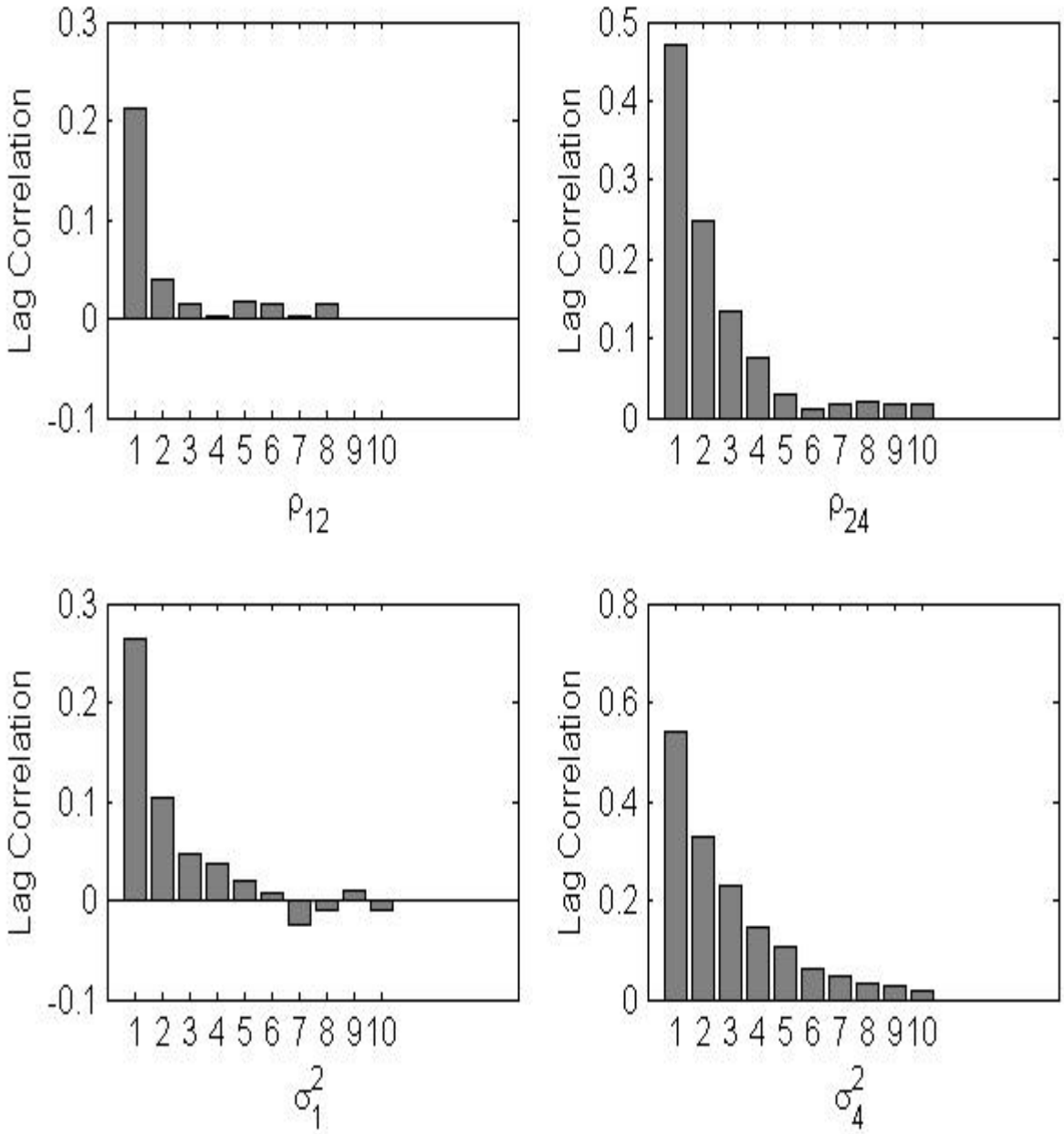

Table 2: True Values and Posterior Estimates of the Parameters

Variable	True Value	Posterior Estimates E(·|y)	Posterior Estimates Std(·|y)
Regression Parameters			
α_1	0.64	0.6323	0.0066
γ_{11}	0.35	0.3547	0.0039
γ_{12}	0.39	0.3852	0.0028
γ_{13}	-0.53	-0.5297	0.0019
γ_{14}	-0.21	-0.2103	0.0031
β_1	-0.49	-0.494	0.0046
α_2	0.93	0.9311	0.0049
γ_{21}	0.39	0.3852	0.0028
γ_{22}	0.3	0.3023	0.0035
γ_{23}	0.2	0.1997	0.0018
γ_{24}	-0.89	-0.8872	0.0029
β_2	0.25	0.2515	0.0041
α_3	-0.12	-0.1213	0.0031
γ_{31}	-0.53	-0.5297	0.0019
γ_{32}	0.2	0.1997	0.0018
γ_{33}	0.1	0.0992	0.0021
γ_{34}	0.23	0.2308	0.0019
β_3	0.34	0.3404	0.0024
α_4	-0.45	-0.4421	0.0063
γ_{41}	-0.21	-0.2103	0.0031
γ_{42}	-0.89	-0.8872	0.0029
γ_{43}	0.23	0.2308	0.0019
γ_{44}	0.87	0.8668	0.0041
β_4	-0.1	-0.098	0.0048
Covariance Matrix Parameters			
ρ_{12}	-0.45	-0.4543	0.0079
ρ_{23}	-0.2	-0.1997	0.0095
ρ_{13}	0.5	0.5034	0.0075
ρ_{14}	-0.35	-0.3405	0.0088
ρ_{24}	0.4	0.3967	0.0085
ρ_{34}	-0.5	-0.4899	0.0076
σ_1^2	0.5	0.505	0.0071
σ_2^2	0.3	0.3038	0.0043
σ_3^2	0.1	0.1008	0.0014
σ_4^2	0.6	0.5982	0.0085

item, whether the price was paid with a deal or not and the coupon value used if any.

The 2006 Homescan data include information from over 37,000 households, although only 7,534 households reported purchases of both random-weight and UPC coded food items. Of these, 7,415 households reported purchases for at least 10 months in 2006. Our final sample comes from the household panel and consists of 7,081 households that had expenditures on ready-to-eat and hot cereals at some time during the year.

We matched the household purchases with the household demographic data. The household characteristics include household size, income, age of household head, education and employment of female and male heads, marital status, race, presence of children and region of residence.

4.2 Whole Grains Identification

We constructed a dataset for purchases of four cereal types: whole grain ready-to-eat, non-whole grain ready-to-eat, whole grain hot, and non-whole grain hot cereals. Although the 2005 Dietary Guidelines recommend that Americans eat three or more one-ounce-equivalent servings of whole grains per day, the government offers no straightforward way for consumers to identify whole grain products, and guidance to the industry on labeling is still not mandated by the Food and Drug Administration. Manufacturers have begun to label their products on whole grain content and the Whole Grains Council provides an approved stamp to indicate products that are good sources of whole grain. ACNielsen provided information on the grain type of some products reported in the HomeScan files. We used these three sources to identify cereals as whole grain and non-whole grain: the Whole Grains Council listing; manufacturers' sites; and the ACNielsen indicator of grain content.

Where information on whole grain content was lacking from the Whole Grain Council, we verified manufacturers' websites and specifically checked if the product was claimed as a whole grain or contained whole grain as a first ingredient. In most cases we were able to identify whole grain products. For example, all General Mills ready-to-eat cereals carry a whole grain claim and listed whole grain as a first ingredient. Many websites had information on ingredients. In some cases, when we were not able to find a manufacturer's whole grain claim, we identified cereals as whole grain if the first ingredient listed was whole grain. Again we found some discrepancies in whole grain coding, but resolved them based on evidence from similar products.

Table 3 shows the total number of UPC's by cereal type in our data set and number and percent of cereals identified as whole grain from the three sources: scanner data "grain type" variable, Whole Grain Council and manufacturer's claim. As indicated in the table, we considered 3810 unique UPC product types; most were in the ready-to-eat cereal category. Included in the data were UPC codes for a large number of private label cereals. Private labels represent 61%, and 68% of total UPC's of ready-to-eat cereals and hot cereals, respectively. Without a manufacturer site, we needed to assign these products to whole grain and non-whole grain product groups. We developed two approaches to classification. In the first, we coded cereals as whole grain if they (a) carried the Whole Grain Council stamp or (b) were identified as a whole grain product by the manufacturer. The remaining products were coded as non-whole grain. In the second approach, we coded products as whole grain if they (a) carried a Whole Grain Stamp, or (b) were identified as a whole grain product by the manufacturer, and the remaining products, including the private labels, were assigned to whole grain if the majority of the observations in the grain type variable were identified as whole grain.

That is, if the private label hot cereal indicated the grain type was "rolled oats", then the private label hot cereal was classified as "whole grain".

Table 3: Cereals Identified as Whole Grain from Different Sources

	Total UPC's	Manufacturer		WG Council		By Grain Type	
	N	N	%	N	%	N	%
Ready-to-eat	2850	514	18.0	198	6.9	603	21.2
Hot Cereal	960	212	22.1	60	6.3	633	65.9
All	3810						

The two resulting classifications are shown in Table 4. As we can see, there are substantial differences in the number of whole grain UPC's identified by the two classifications. From the total of 2,850 different UPCs available for ready-to-eat cereals, only 18 % is identified as whole grain by classification 1 and almost double of this amount is identified as whole grain by classification 2. With respect to hot cereals, 91% of all UPCs available are identified as whole grain by classification 2, compared to only 22% by classification 1. Compared to classification 1, which assigns all private labels to non-whole grain group, classification 2 seems more reasonable. Although some concerns may be raised regarding the sensitivity of the analysis to the classifications used, it is clear that estimating a demand system using classification 1 can could lead to unreliable results.

Table 4: Classification of Cereals into Whole Grain

	Total UPC's	Classification 1		Classification 2	
	N	N	%	N	%
Ready-to-eat	2850	519	18.2	938	32.9
Hot Cereal	960	212	22.1	877	91.4
All	3810				

4.3 Variables and Descriptive Statistics

The data include repeated expenditures and quantities for each purchased item. The price of each commodity was calculated as the unit value, defined as the aggregated household expenditure for the product divided by quantity purchased in ounces (reported for the year). The household's expenditure was calculated by subtracting the value of any coupons used during the purchase from the amount paid. We also calculated average regional prices. The dataset provides information on 52 Scantrack markets and rural areas. We derived average prices for all four commodities by 52 Scantrack markets and rural areas. For households not purchasing a particular product, we replaced missing prices with the average prices (unit values) based on prices paid by the purchasing households for the household's corresponding market area.

Table 5 presents purchase frequencies, mean expenditure shares, mean expenditures, quantities purchased and unit values for the purchasing households for the commodities used in the analysis. Whole grain ready-to-eat cereal was consumed by the majority of the households and also had the highest mean expenditure and expenditure share among different types of cereals. Ready-to-eat non-whole grain cereal was next most frequently purchased by the households in our sample.

Table 5: Distribution of Purchasing Households and Sample Mean Values of Selected Variables

Product Category	No. of Hhlds	% of Hhlds	Mean Expenditure Share	Mean Quantity (ounces)	Mean Expenditure ($)	Mean Unit Value ($/ounce)
Sample	7081	100.0				
Ready-to-eat WG	6382	90.1	0.48	255.80	39.43	0.16
Ready-to-eat Non-WG	5960	84.2	0.34	183.08	27.60	0.16
Hot Cereal WG	4414	62.3	0.14	99.00	12.40	0.14
Hot Cereal Non-WG	1922	27.1	0.04	84.42	6.47	0.10

Table 6 presents the definitions of the variables used in the analysis along with the means and standard deviations of the variables for the whole sample. The average household income was $59,270. The average household size was 2.34, 23 percent of the sample were households with children, and 59 percent were married couple households. For the analysis reported in this paper, the estimates were unweighted.

Table 6: Definition of Variables, Sample Mean Values and Standard Deviations

Variable	Definition	Mean	Std.
N	Number of households	7081.00	
Income/$1000	Household income/$1000	59.27	39.02
Household size	Household size	2.34	1.29
Age of Head<30	1 if household heads age is under 30	0.01	0.09
30≤Age of Head ≤49	1 if household heads age is between 30&49	0.31	0.46
50≤Age of Head≤64	1 if female heads age is between 50&64	0.40	0.49
65≤Age of Head	1 if female heads age is 65 and older	0.28	0.45
Presence of children	1 if household has children	0.23	0.42
Male head employed	1 if the male head is employed	0.66	0.47
Female head employed	1 if the female head is employed	0.59	0.49
≤ High school (male)	1 if the male heads education is high school	0.27	0.44
Some college (male)	1 if the male heads education is some college	0.31	0.46
College + (male)	1 if the male heads education is college	0.43	0.49
≤ High school (female)	1 if female heads education is high school	0.27	0.44
Some college (female)	1 if the female heads education is some college	0.31	0.46
College (female)	1 if female heads education is college	0.41	0.49
Married	1 if married	0.59	0.49
White	1 if race is white	0.77	0.42
Black	1 if the race is black	0.13	0.34
Other	1 if race is other	0.10	0.30
Hispanic	1 if Hispanic	0.07	0.26
East	1 if the household lives in the East region	0.22	0.42
Central	1 if the household lives in the Central region	0.17	0.37
South	1 if the household lives in the South region	0.38	0.49
West	1 if the household lives in the West region	0.23	0.42
Urban	1 if the household lives in urban area	0.87	0.34

Table 7 presents the means and standard deviations for the variables used in the model for the four commodities. As indicated in Table 7, not much difference exists

among the mean values of the variables across product categories, except for some variables of households purchasing non-whole grain hot cereals. These households were more likely to have lower income, be over the age of 65 and be married compared to the other three groups.

Table 7: Variables and Sample Mean Values (N=7081)

	Ready-to-Eat (n=6875)		Hot Cereal (N=5031)	
Variable	WG	Non-WG	WG	Non-WG
N	6382	5960	4414	1922
Income/$1000	60.19	59.43	60.82	55.76
Household size	2.40	2.45	2.43	2.47
Age of Head<30	0.01	0.01	0.01	0.01
30≤Age of Head ≤49	0.32	0.33	0.30	0.30
50≤Age of Head≤64	0.40	0.39	0.39	0.36
65≤Age of Head	0.28	0.27	0.30	0.33
Presence of Children	0.24	0.25	0.24	0.25
Male Head Employed	0.67	0.67	0.65	0.62
Female Head Employed	0.59	0.59	0.58	0.54
≤ High School (male)	0.27	0.28	0.26	0.29
Some College (male)	0.31	0.31	0.31	0.30
College + (male)	0.43	0.41	0.42	0.41
≤ High School (female)	0.27	0.28	0.27	0.29
Some College (female)	0.32	0.32	0.33	0.32
College + (female)	0.41	0.40	0.40	0.39
Married	0.61	0.62	0.63	0.65
White	0.78	0.77	0.78	0.79
Black	0.13	0.13	0.13	0.13
Other	0.10	0.10	0.10	0.08
Hispanic	0.08	0.08	0.08	0.08
East	0.23	0.23	0.23	0.21
Central	0.17	0.17	0.16	0.20
South	0.38	0.38	0.38	0.38
West	0.23	0.22	0.23	0.21
Urban	0.87	0.87	0.87	0.83

5 Empirical Results

A system of four equations was estimated using data based on classification 2 in Table 4 (the classification that assigns whole grain values to private label items). We fit our model using the algorithm specified in previous section. We ran our posterior simulator for 100,000 iterations and discarded the first 30,000 as the burn-in. For our prior hyper-parameters, we set $\mu_{\boldsymbol{\theta_2}}$ equal to a zero vector of the dimension $(k-J)\times 1$, $\boldsymbol{V}_{\boldsymbol{\theta_2}}$ and $\underline{\mathbf{A}}$ to identity matrices of the appropriate dimensions and $\underline{\nu}=5$.

Tables 8 and 9 present the posterior means, posterior standard deviations and probabilities of being positive for the demographic, price and expenditure related parameters for whole grain and non-whole grain ready-to-eat and hot cereals, respectively. We find that larger households are less likely to consume either type of whole-grain cereals and more likely to consume non-whole grain cereals, both ready-to-eat and hot. Households with higher income tend to consume more whole grain ready-to-eat and non-whole grain hot cereals and less non-whole grain ready-to-eat and whole grain hot cereals. Households with children present are more likely to consume both types of ready-to-eat cereals and less likely to consume both types of hot cereals. There are some race/ethnic differences. Ready-to-eat cereal is a prevalent food in the diets of Americans, especially children (Song and et al. 2006).

Estimated parameters were used to calculate price and cereal expenditure elasticities in order to examine the responsiveness of the consumers to economic incentives (Table 10). The uncompensated and compensated own-price elastiticities are all negative, as expected for normal goods for which demand responds negatively to increases in prices. The uncompensated (Marshallian) price elasticities include an income effect as well as price effect.

Table 8: Ready-to-Eat Cereals: Posterior Means and Probabilities of Being Positive

	Ready-to-Eat					
	WG			Non-WG		
Variable	E(·|y)	Std(·|y)	Pr(· > 0|y)	E(·|y)	Std(·|y)	Pr(· > 0|y)
Demographic Characteristics						
Intercept	0.2333	0.0041	1	0.3559	0.0038	1
Income/ $1000	0.0006	0	1	-0.0005	0	0
Household size	-0.0201	0	0	0.0266	0	1
Age of Head<30	0.0506	0.0003	1	0.0641	0.0003	1
30 ≤ Age of Head ≤ 49	-0.0244	0.0001	0	0.0629	0.0001	1
50 ≤ Age of Head ≤ 64	-0.0005	0.0001	0	0.0126	0.0001	1
Presence of Children	0.0076	0.0001	1	0.0051	0.0001	1
Male Head Employed	0.0126	0.0001	1	0.001	0.0001	0.8
Female Head Employed	-0.0084	0.0001	0	0.0087	0.0001	1
≤ High School (male)	-0.0245	0.0001	0	0.0354	0.0001	1
Some College (male)	-0.0137	0.0001	0	0.0176	0.0001	1
≤ High School(female)	-0.0037	0.0001	0	0.0171	0.0001	1
Some College (female)	-0.0063	0.0001	0	0.011	0.0001	1
Married	0.0105	0.0001	1	-0.0093	0.0001	0
White	0.0547	0.0002	1	-0.0389	0.0002	0
Black	-0.0262	0.0003	0	0.0188	0.0003	1
Hispanic	-0.0197	0.0003	0	0.0171	0.0003	1
East	0.0229	0.0002	1	-0.0022	0.0001	0
Central	-0.0015	0.0003	0	0.0092	0.0002	1
South	-0.0074	0.0003	0	0.0145	0.0002	1
Urban	0.0164	0.0002	1	-0.0142	0.0002	0
Price Coefficients						
RTE WG	0.0909	0.0008	1	-0.0536	0.0006	0
RTE NWG	-0.0536	0.0006	0	0.0909	0.0004	1
Hot WG	0.0044	0.0004	1	-0.0066	0.0003	0
Hot Non-WG	-0.0417	0.0017	0	-0.0307	0.0013	0
Total Expenditure						
Expenditure	0.0388	0.0001	1	-0.0107	0.0001	0

The values of uncompensated own-price elasticities range from -0.89 for non-whole grain hot cereals to -0.44 for non-whole grain ready-to-eat. All are price inelastic, with largest (absolute) values being for whole grain cereals.

The mean unit prices for hot cereals reported in Table 5, especially for non-whole

Table 9: Hot Cereals: Posterior Means and Probabilities of Being Positive

	Hot Cereal					
	WG			Non-WG		
Variable	E(·|y)	Std(·|y)	Pr(· > 0|y)	E(·|y)	Std(·|y)	Pr(· > 0|y)
Demographic Characteristics						
Intercept	0.4335	0.0035	1	-0.0227	0.0114	0
Income/ $1000	-0.0001	0	0	0.001	0	1
Household size	-0.0079	0	0	0.0014	0	1
Age of Head<30	-0.1059	0.0003	0	-0.0088	0.0005	0
30 ≤ Age of Head ≤ 49	-0.0346	0.0001	0	-0.0039	0.0001	0
50 ≤ Age of Head ≤ 64	-0.0107	0.0001	0	-0.0014	0.0001	0
Presence of Children	-0.0102	0.0001	0	-0.0025	0	0
Male Head Employed	-0.0136	0.0001	0	0.0009	0.0002	1
Female Head Employed	0.0027	0.0001	1	-0.003	0.0003	0
≤ High School (male)	-0.0124	0.0001	0	0.0015	0.0001	1
Some College (male)	-0.0075	0.0001	0	0.0037	0.0002	1
≤ High School(female)	-0.0102	0.0001	0	-0.0032	0	0
Some College (female)	-0.005	0.0001	0	0.0002	0.0001	1
Married	-0.0017	0	0	0.0006	0	1
White	-0.0291	0.0002	0	0.0133	0.0005	1
Black	-0.0152	0.0003	0	0.0226	0.0009	1
Hispanic	-0.0137	0.0003	0	0.0163	0.0008	1
East	-0.0262	0.0001	0	0.0056	0.0004	1
Central	-0.0218	0.0002	0	0.0141	0.0006	1
South	-0.0208	0.0002	0	0.0137	0.0007	1
Urban	0.0072	0.0001	1	-0.0095	0.0005	0
Price Coefficients						
Hot WG	0.0235	0.0002	1	-0.0213	0.001	0
Hot Non-WG	-0.0213	0.001	0	0.0937	0.004	1
Total Expenditure						
Expenditure	-0.024	0.0001	0	-0.0042	0.0003	0

grain, were relatively smaller compared to mean unit prices for both ready-to-eat cereals and whole-grain hot cereals. Most of the (Hicksian) cross-price elasticities are positive indicating substitutability among the cereal types. Results indicate that most of the cross-price elasticities are small; the largest one is between the ready-to-eat and hot whole grain cereals. Relatively lower values (in absolute terms) for the cross-price effects

Table 10: Estimated Demand Elasticities

	Ready-to-Eat				Hot Cereals			
	WG		Non-WG		WG		Non-WG	
Variable	E(·|y)	Pr(· > 0|y)	E(·|y)	Pr(· > 0|y)	E(·|y)	Pr(· > 0|y)	E(·|y)	Pr(· > 0|y)
Marshallian Elasticity								
RTE WG	-0.7764	0	-0.166	0	-0.0161	0	-0.1505	0
RTE Non-WG	-0.2879	0	-0.4415	0	-0.016	0	-0.1894	0
Hot WG	0.0485	1	-0.0062	0	-0.8926	0	-0.066	0
Hot Non-WG	-0.2363	0	-0.1761	0	-0.1082	0	-0.4425	0
Hicksian Elasticity								
RTE WG	-0.3842	0	0.2262	1	0.3762	1	0.2418	1
RTE Non-WG	-0.1309	0	-0.2844	0	0.1411	1	-0.0324	0
Hot WG	0.3133	1	0.2586	1	-0.6278	0	0.1988	1
Hot Non-WG	-0.0504	0	0.0099	1	0.0778	1	-0.2566	0
Expenditure Elasticity								
	1.109	1.000	0.9348	1.000	0.9163	1.000	0.9821	1

indicate that consumers are more responsive to own-price rather than prices of the other goods.

The total expenditure elasticities do not vary widely in the magnitude. The total expenditure elasticity is slightly above unity for the whole grain ready-to-eat cereals suggesting that as the expenditure on cereals increases households will allocate proportionally more on whole-grain ready-to-eat cereals and less on other cereals.

6 Discussion and Conclusion

This paper describes a procedure for estimating a censored AIDS model using Bayesian methods. ACNielsen 2006 scanner data are used in estimating the demand for breakfast cereals. We disaggregate the cereals by grain type and by type of cereal and estimate the system of four equations. Within the cereal groups we find demand for all four cereals to be price inelastic. Demand for whole grain hot cereals (which includes rolled oats) is the most sensitive to price changes. Cross price elasticities indicate consumers substitute among the four types, although the cross-price substitution effects (elasticities) are small.

Results of this research can be sensitive to this classification, since more that 50 percent of cereals in our data carry private labels. Also, although the observed shares for the four products we analyzed do add-up to unity, by construction, the estimation method we used does not account for the adding-up to unity of the latent expenditure shares. This is a complex estimation problem and further work is needed to specify a model that imposes an adding-up to unity restriction on the latent expenditure shares.

Understanding consumer willingness to substitute between "healthy" and "not so healthy" products is critical to designing effective health and food policies and messages based on consumer behaviors. Often product groups of interest are not widely consumed

in the population, and therefore, there are issues of zero expenditure shares. We propose a method for addressing this methodological challenge. We find that among grain products, consumers do make substitutions although only to a limited extent. Messages and nutrition guidance designed to encourage whole grains consumption may encourage a greater willingness to substitute. Improved food labeling would allow consumers to identify whole grains more easily. And, as manufacturers respond with a wider selection of whole grain products, we would expect greater consumer awareness of substitute products. Although our evidence indicates relatively small substitution effects, the cross product substitution effects are positive and suggest opportunity to exploit a willingness by consumers to switch to more healthful product types.

7 Acknowledgements

We acknowledge USDA/Economic Research Service for providing partial financial support and access to the ACNielsen Homescan data.

References

Albert, J., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88, 669-679.

Amemiya, T. 1974. Multivariate regression and simultaneous equation models when the dependent variables are truncated normal. *Econometrica* 42: 999-1012.

Buzby, J., H. Farah and G. Vocke. 2005. Will 2005 Be the Year of the Whole Grain? Amber Waves, June.

Deaton A. and J. Muellbauer. 1980. An Almost Ideal Demand System. *American Economic Review* 70:312-326.

Dong, D., B.W. Gould and H.M. Kaiser. 2004. Food demand in Mexico: An application of the Amemiya-Tobin approach to the estimation of a censored food system. *American Journal of Agricultural Economics* 86:1094-1107.

Fung, T.T., F.B. Hu, M.A. Pereria, S. Liu, M.J. Stampfer G.A. Colditz, and W.C. Willett. 2002. Whole-grain intake and the risk of type 2 diabetes: a prospective study in men. *American Journal of Clinical Nutrition* 76:535-40.

Geweke, J. 1991. Efficient simulation from the multivariate Normal and Student-t distributions subject to linear constraints, *in Computer Science and Statistics: Proceedings of the 23rd Symposium on the Interface* (ed. E. Keramidas) 571-578.

Golan A., Perloff J. and E. Shen. 2001. Estimating a demand system with non-negativity constraints: Mexican meat demand. *Review of Economics and Statistics* 83:541-550.

Griffiths, E., C.J. O'Donnell, and A. Tan Cruz. 2000. Imposing regularity conditions on a system of cost and factor share equations. *The Australian Journal of Agricultural and Resource Economics* 44:101-127.

Hallfrisch, J., D.J. Scholfield, and K. M. Behall. 2003. Blood pressure reduced by whole grain diet containing barley or whole wheat and brown rice in moderately hypercholesterolemic men *Nutrition Research* 23 1631-1642.

Huang, C.J., F.A. Sloan, and K.W. Adamache. 1987. Estimation of seemingly unrelated Tobit regressions via the EM algorithm. *Journal of Business and Economic Statistics* 5, 425430.

Jacobs DR, Marquart L, Slavin J, Kushi LH. 1998. Whole grain intake and cancer: An expanded review and meta-analysis. *Nutrition and Cancer* 30:85-96.

Koop, G., D.J. Poirier and J.L. Tobias. 2007. *Bayesian Econometric Methods*, Cambridge, Cambridge University Press.

Lee, L.F. and M. Pitt. 1986. Microeconometric demand systems with binding non-negativity constraints: A dual approach. *Econometrica* 55:1237-42.

Lin, B.H. and S.T.Yen 2007. The U.S. grain consumption landscape: who eats grain, in what form, where and how much? Economic Research Report-50. U.S. Department of Agriculture, Economic Research Service.

Liu S., J.E. Manson, M.J. Stampfer, F.B. Hu, E. Giovannucci, G.A. Colditz, C.H. Hennekens, and W.C. Willett. 2000. "A prospective study of whole-grain intake and risk of type 2 diabetes mellitus in US women. *American Journal of Public Health* 90:1409-15.

Liu S., M.J. Stampfer, F.B. Hu, E. Giovannucci, E. Rimm, J.E. Manson, C.H. Hennekens, and W.C. Willett. 1999. Whole-grain consumption and risk of coronary heart disease: results from the Nurses' Health Study. *American Journal of Clinical Nutrition* 70:412-9.

Lupton, J.R., M.C. Robinson, and J.L. Morin. 1994. Cholesterol-lowering effect of barley bran flour and oil. *Journal of American Dietetic Association* 94:65-70.

Mancino, L. and J. Buzby. 2005. Americans' Whole-Grain Consumption Below Guidelines. Amber Waves. April.
http://www.ers.usda.gov/Amberwaves/April05/Findings/WholeGrainConsumption.htm

Perali F., and J.P. Chavas. 2000. "Estimation of censored demand equations from large cross-section data. *American Journal of Agricultural Economics* 82:1022-1037.

Science Daily. 2007. Whole-grain breakfast cereal associated with reduced heart failure risk.
http://www.sciencedaily.com/releases/207/03/070302171407.htm

Shonkwiler, J.S. and S.T. Yen. 1999. Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics* 81:972-82.

Song,W.O., O.K. Chun, J. Kerver, S. Cho, C.E.Chung and S.J. Chung. 2006. *Journal of American Dietetic Association* 106:1783-1789.

Tobin, J. 1958. Estimation of relationships for limited dependent variables. *Econometrica* 26:24-36.

Truswell, A.S. 2002. Cereal grains and coronary heart disease. *European Journal of Clinical Nutrition* 56:1-14.

U.S. Department of Health and Human Services and U.S. Department of Agriculture. 2005. Dietary Guidelines for Americans. Available at:
http://www.health.gov/dietaryguidelines/dga2005/document/pdf/DGA2005.pdf

U.S. Department of Health and Human Services and U.S. Department of Agriculture. Public comment period begins on report of dietary guidelines advisory committee. Available at:
http://www.hhs.gov/new/press/2004pres/20040827.html

U. S. Food and Drug Administration. Center for Food Safety and Applied Nutrition. Of-

fice of Nutritional Products, Labeling and Dietary Supplements. 2008. Guidance for industry: A food labeling guide. Appendix C.

Wales, T.J. and A.D. Woodland. 1983. Estimation of consumer demand system with binding non-negativity constraints. *Journal of Econometrics*, 21: 263-85.

Whole Grain Council, 2008. Whole Grain Statistics. Available at: http://www.wholegrainscouncil.org/newsroom/whole-grain-statistics

Yen S.T. 2005. A multivariate sample-selection model: estimating cigarette and alcohol demands with zero observations. *American Journal of Agricultural Economics*, 87:453-466

Yen S.T., Kan K., Su S-J. 2002. Household demand for fats and oils: two-step estimation of a censored demand system. *Applied Economics*, 14:1799-1806.

Yen, S., Lin, B. H. and Smallwood, D. M.. 2003. Quasi- and simulated-likelihood approaches to censored demand system: food consumption by food stamp recipients in the United States. *American Journal of Agricultural Economics*, 85:458-478.

Yen S.T., Roe T.L. 1989. Estimation of a two-level demand system with limited dependent variables. *American Journal of Agricultural Economics*, 71:85-98.

Zellner A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57:348-368.