

**Bayesian Estimation of The Impacts of Food Safety Information on  
Household Demand for Meat and Poultry**

by

**Mykel R. Taylor**

School of Economic Sciences  
Washington State University  
P.O. Box 646210  
Pullman, WA 99164-6210  
E-mail: m\_taylor@wsu.edu

**Daniel Phaneuf**

Department of Agricultural and Resource Economics  
North Carolina State University  
Box 8109  
Raleigh, NC 27695-8109  
E-mail: dan\_phaneuf@ncsu.edu

*Selected Paper prepared for presentation at the Agricultural & Applied Economics  
Association's 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, WI, July 26-28, 2009.*

*Copyright 2009 by Taylor and Phaneuf. All rights reserved. Readers may make verbatim copies of  
this document for non-commercial purposes by any means, provided this copyright notice appears on  
all such copies.*

Many factors can influence consumer purchasing habits, including information about product safety. Concerns about food safety are likely to be influenced by idiosyncratic experiences such as suffering from a foodborne illness or receiving medical warnings from a physician regarding susceptibility to bacterial pathogens.<sup>1</sup> General media information on the safety of meat and poultry might also affect purchase decisions. This is particularly plausible when large scale food safety events occur and media coverage of contaminated meat or poultry products is heightened. While idiosyncratic experiences are difficult to measure, the amount of food safety information available to consumers in the press can be quantified.

The objective of this study is to investigate if the quantity of publicly available food safety information impacts consumers' decisions to purchase fresh meat and poultry. A media index measuring the number of articles containing food safety information on beef, pork, or poultry published in U.S. regional newspapers is used as a proxy for food safety information available to consumers. The media index is a broad measure in that it includes reporting on domestic recall events as well as international issues, commentary on food contamination prevention, and other food safety-related topics. Commodity- and region-specific, monthly parameters are constructed using the media index and a discrete-continuous choice tobit model is estimated to measure the impact of food safety information on purchase behavior. Results from the study will provide insight into households' propensity to avoid or change their consumption of a commodity when faced with food safety concerns.

## **Literature Review**

Previous research on consumer responses to food safety information has employed various measures of media coverage to infer its effect on food demand. Dahlgran and Fairchild

---

<sup>1</sup> An example of a food safety warning from a physician would be providing information to pregnant women on the increased risks of miscarriage due to listeria contamination.

(2002) studied the effect of adverse media coverage from salmonella contamination on the demand for chicken. Their model incorporated adverse media publicity from T.V. and print as a form of negative advertising, where publicity included both the number of stories aired and the percent of population exposed to the coverage. Weekly market-level data on quantity and prices of chicken were used to allow measurement of short-run effects on the price of chicken. Their results did indicate a negative demand response to adverse media, however, the effect died out in a matter of weeks. Burton and Young (1996) analyzed the effects of bovine spongiform encephalopathy (BSE) on meat demand in Great Britain using media indices incorporated into a dynamic AIDS model. The analysis used quarterly data on quantity and expenditures for beef, lamb, pork, and poultry. The study used an index of media coverage and showed that BSE publicity had both significant short-run and long-run effects on consumer expenditures on beef and among the other meats.

A recent study by Piggott and Marsh (2004) analyzed the impact of food safety information on demand for beef, pork, and poultry using aggregate data on quarterly U.S. per capita disappearance of meat. They developed a theoretical model that incorporated meat quality into the demand for meat. The framework also explicitly considered both own- and cross-product effects from quality on the quantity demanded. Meat quality, in their model, was inversely related to the occurrence of food safety information in the media. The media index for food safety information measured bundles of contaminants reported individually for beef, pork, and poultry. Their findings indicated that effects of food safety information on meat demand were statistically significant, but with no lagged effect implying a relatively small economic impact.

Each of these studies used aggregate data to estimate meat and/or poultry demand equations that quantify the effect of food safety information on purchases. This approach has

shown that media information matters at the aggregate level, however it does not allow assessment of the likelihood that individual households will avoid purchasing meat and poultry products all together in response to food safety information. Examining both marginal and discrete avoidance behavior at the disaggregate level (i.e., what mix of products households buy on a given purchase occasion) can provide additional insight into consumer demand for meat and poultry products under different food safety information environments.

## **Data**

Monthly data from the time period January 1998 to December 2005 is used to analyze the effects of food safety information on U.S. household demand for meat and poultry. The data for this study come from two sources. Data on household purchases of meat and poultry were obtained from the Nielsen Homescan panel. These panel data also contain information on several demographic characteristics of the participating households. The data used to describe food safety information were obtained from searches of newspapers using the Lexis-Nexis academic search engine.

The Nielsen Homescan panel is a nationwide survey of households and their retail food purchases. Households record purchase data by scanning the universal product codes (UPCs) of the items they purchase. Each item is recorded by a scanning device at home after every shopping trip. The purchase data are subsequently uploaded electronically to Nielsen's database. Data include detailed product information, date of purchase, total quantity, total expenditure, and the value of any coupons used for every item purchased. Not all food products are marked with a UPC code. Unmarked items are referred to as random-weight products and include foods such as fresh meat and poultry or fresh fruits and vegetables. Random weight items are recorded by using a code book provided by Nielsen that contains product descriptions and unique codes that

can be scanned by the individual. Both random-weight and UPC coded products are used in the analysis.

The products of interest for this study are fresh and frozen beef and veal, pork, chicken, and turkey. These groups do not include any processed products because it becomes difficult to determine the extent of processing and the value added to the final price from processing.<sup>2</sup> All the fresh products used in the proposed demand analysis are random-weight items and the frozen products are marked by a UPC code. Each observation is a separate product purchase and includes the total quantity purchased in pounds, the total amount spent on the item in dollars, a product description (e.g. ground beef-bulk, rib eye steak, whole chicken), and the date of purchase. Prices per unit of product were subsequently calculated by dividing total expenditure by total quantity for each individual meat or poultry purchase.

One advantage of working with daily purchase data is the flexibility to choose the frequency of observation. The choice of periodicity is driven primarily by the level of censoring in the data. If purchases were aggregated to a weekly level, the amount of censoring in this dataset is very large. Quarterly data greatly reduces the amount of censoring for all commodities, but that level of periodicity could mask possible short run food safety effects. Therefore, a compromise of a monthly periodicity was chosen for the empirical analysis. Approximately 4.70% of the households did not purchase any meat or poultry products in a given year. These households were removed from the panel, leaving 62,136 households across all eight sample years.

The Nielsen Homescan panel is a stratified random sample that was selected based on both geographic and demographic targets. The dataset used in this study is an unbalanced panel

---

<sup>2</sup> Examples of processed meat and poultry products include luncheon meats, frozen dinners, or soups that contain meat or poultry.

in that not all households participated for all sample years. However, the distributions of the demographic and geographic characteristics of the households within a sample year do not vary noticeably from year to year. Summary statistics of the household demographic variables are listed in table 1.

As mentioned previously, prices per unit of each meat and poultry product were calculated by dividing total expenditure by total quantity. This results in retail prices being available only for the households that actually made purchases. For the households that chose not to purchase a product in a given month, the price they faced for that product is not recorded. Therefore, the missing prices must be imputed for households without positive purchases in order to have a complete dataset for estimation purposes. Following Cox and Wohlgenant (1986), household income is used to capture hypothesized increases in quality that may be demanded from increased income. A variable for household size is used to account for economies of size in purchasing meat and poultry products. Quadratic terms for both income and household size are also included in the regression. Other demographic variables were considered for the price equations; however, the coefficients were not statistically different from zero for most of the goods.

The final specification of the linear price regression is as follows:

$$p_{im} = \alpha \bar{p}_{it} + \gamma_r \mathbf{r}_n + \delta u_n + \eta i_n + \kappa i_n^2 + \tau s_n + \rho s_n^2 + \varepsilon_{it} , \quad (1)$$

where  $p_{im}$  is the observed price of good  $i$  in month  $t$  for consuming household  $n$ ,  $\bar{p}_{it}$  is the sample average monthly price for good  $i$  in month  $t$ ,  $\mathbf{r}_n$  is a vector of binary variables indicating the region in which the household is located,  $u_n$  is a binary variable indicating if the household is located in an urban area,  $i_n$  is household income,  $i_n^2$  is household income squared,  $s_n$  is the size

of household,  $s_n^2$  is the squared size of household,  $\varepsilon_{it}$  is an iid error term, and  $\alpha, \gamma_r, \delta, \eta, \kappa, \tau$ , and  $\rho$  are the corresponding coefficients to be estimated.<sup>3</sup> The regression is estimated without a constant term so that all the regional binary variables can be included and standard errors are estimated using the robust sandwich estimator (Huber, 1967; White, 1980).

The regression coefficients for each good were subsequently used to predict prices for the non-consuming households. Predicted prices were obtained by using the sample monthly average prices and the geographic and demographic characteristics of the non-consuming households. These predicted prices replace the zeros to provide a complete series of prices for subsequent demand analysis.

The grouping of purchases into various beef, pork, and poultry products of similar characteristics and average prices is intended to minimize the amount of quality and price variation that occurs when the daily purchases are aggregated to a monthly level. However, the number of equations that must be estimated is still relatively large (five beef, four pork, and six poultry groups), so the products are aggregated to the commodity level for estimation purposes. While aggregation is useful for estimation, it can mask variation in product prices and quality, making explicit consideration of this variation within aggregate commodities critical.

In order to account for the within-species price and quality variation that exists when purchases were aggregated, a Törnqvist (1936) price index was used. The expenditure share-weighted geometric price index defined as follows:

$$P_{nt}^B = \prod_{i=1}^G P_{\text{int}}^{w_i} \quad , \quad (2)$$

---

<sup>3</sup> Total household income is recorded as an interval in this dataset. Therefore, the midpoint of the interval is the value used in the price regression. To calculate the midpoint of the highest income range, an upper bound of \$150,000 was used.

where  $p_{nt}^B$  is the index price of beef for household  $n$  in month  $t$ ,  $p_{int}$  is the retail price of beef group  $i$  faced by the household  $n$  in month  $t$ ,  $w_i$  is the beef group  $i$  share of total household expenditures on all groups of beef, and  $G$  is the number of groups specified for beef. The expenditure share is calculated as follows:

$$w_i = \frac{\bar{p}_i \bar{x}_i}{\sum_{j=1}^G \bar{p}_j \bar{x}_j}, \quad (3)$$

where  $\bar{p}_i$  is the average price of beef group  $i$  across the entire sample period and  $\bar{x}_i$  is the average quantity purchased of beef group  $i$  across the entire sample period.<sup>4</sup> For beef, there are five subgroups with group 1 referring to ground beef, group 2 to roasts, group 3 to steaks, group 4 to frozen beef, and group 5 to other beef. A similar price index was calculated for the pork and poultry aggregates as well, using four groups for pork and six groups for poultry. The summary statistics of the price and quantity indices are listed in table 2.

Following Piggott and Marsh (2004), food safety is measured using commodity-specific indices of newspaper articles. This specification of commodity-specific media indices allows the cross-commodity effects of food safety information to be explicitly modeled. Relevant articles from six major papers in each of four regions of the United States were found using the Lexis-Nexis search engine. The articles counts gathered from the regional newspaper search were aggregated to create indices that are 30-day rolling averages of the number of newspapers articles published during the previous two weeks.<sup>5</sup> The intuition for this specification of the indices is that each day of the month is a potential purchase occasion and the available and

---

<sup>4</sup> The monthly retail price of each group is the observed group price if the household bought that group in month  $t$ . If the household did not purchase that group, then the predicted group price is used.

<sup>5</sup> The choice of a two week ‘memory’ for the media index is based on investigation of the household purchase data. These data indicate that, on average, fresh meat and poultry products are bought about 2 times per month.



relevant information for each purchase occasion may change as time passes. At the beginning of the month, the articles most likely to impact household purchase decisions are the ones published in the latter half of the previous month. Over the course of the month, however, the most relevant food safety information becomes articles published in the current month. The rolling average specification captures this change in available information over the 30 day period. Figures 1-3 display the regional media indices for each of the three commodity groups.

### **Demand Model**

The demand model is estimated as a seemingly unrelated regression (SUR) tobit model. There are two reasons for the use of this particular estimator. First, not all households buy all three of the commodities considered in this study every month. If an ordinary least squares (OLS) estimator were used for this analysis, the resulting coefficients would be biased toward zero with the degree of bias increasing as the percentage of censoring increases. The second reason a SUR tobit model was chosen is due to the possible correlation that exists between the errors of the beef, pork, and poultry demand equations. These three commodities are likely to be substitutes and consumer's decisions of which product to buy are potentially affected by characteristics of the others. The use of a system estimator such as SUR will explicitly account for any error correlation that may exist between the three commodities, providing more efficient estimates than single equation estimation.

The SUR tobit model is specified with a component error structure (random effects model) to account for the correlation that is likely to exist between observations from the same household. The random effects SUR tobit model is comprised of  $J$  commodities (equations) and  $(N \cdot T)$  outcomes where  $N$  is the number of households and  $T$  is the total number of times all the households appear in the dataset. The model is specified as follows:

$$y_{ijt}^* = \alpha_j + \mathbf{x}_{ijt}\boldsymbol{\beta}_j + \mathbf{c}_i\boldsymbol{\gamma}_j + u_{ij} + \varepsilon_{ijt}, \quad j=1,\dots,J, \quad i=1,\dots,N, \quad t=1,\dots,T_i, \quad (4)$$

$$y_{ijt} = \begin{cases} y_{ijt}^* & \text{if } y_{ijt}^* > 0 \\ 0 & \text{if } y_{ijt}^* \leq 0 \end{cases}, \quad (5)$$

where  $u_{ij}$  is the household- and commodity-specific random error term that does not vary over time,  $u_{ij} \sim iid N(0, \sigma_{u_j}^2)$ ,  $T_i$  is the size of the panel for the  $i^{\text{th}}$  household, and all other terms are as defined above with an additional  $t$  index. In an unbalanced panel dataset like the one used in this study,  $T_i$  will vary over households. The system of equations is stacked over  $J$  commodities and written as:

$$\begin{bmatrix} y_{it1}^* \\ y_{it2}^* \\ \vdots \\ y_{itJ}^* \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_J \end{bmatrix} + \begin{bmatrix} \mathbf{x}_{it1} & 0 & \cdots & 0 \\ 0 & \mathbf{x}_{it2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_{itJ} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_J \end{bmatrix} + \begin{bmatrix} \mathbf{c}_i & 0 & \cdots & 0 \\ 0 & \mathbf{c}_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{c}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\gamma}_J \end{bmatrix} + \begin{bmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{iJ} \end{bmatrix} + \begin{bmatrix} \varepsilon_{it1} \\ \varepsilon_{it2} \\ \vdots \\ \varepsilon_{itJ} \end{bmatrix}, \quad (6)$$

or

$$\mathbf{y}_{it}^* = \boldsymbol{\alpha} + X_{it}\boldsymbol{\beta} + C_i\boldsymbol{\gamma} + \mathbf{u}_i + \boldsymbol{\varepsilon}_{it}, \quad (7)$$

for  $i=1,\dots,N$ ,  $t=1,\dots,T_i$ . Combining the regressor matrices,  $X_{it}$  and  $C_i$ , equation (7) can be rewritten as:

$$\mathbf{y}_{it}^* = W_{it}\boldsymbol{\theta} + \mathbf{u}_i + \boldsymbol{\varepsilon}_{it}, \quad (8)$$

where  $W_{it} = [I_J \ X_{it} \ C_i]$ ,  $\boldsymbol{\theta} = [\boldsymbol{\alpha} \ \boldsymbol{\beta} \ \boldsymbol{\gamma}]$ ,  $I_J$  is a  $(J \times J)$  identity matrix,  $\mathbf{u}_i \sim iid N(0, V)$ , and

$\boldsymbol{\varepsilon}_{it} \sim iid N(0, \Sigma)$  with  $E(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{is}') = 0$  for all  $t \neq s$ . The covariance matrix  $V$  is defined as follows:

$$V = \begin{bmatrix} \sigma_{u_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{u_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{u_J}^2 \end{bmatrix}. \quad (9)$$

The system of equations are further stacked over all households and time periods in the panel and written as:

$$\begin{bmatrix} \mathbf{y}_{11}^* \\ \vdots \\ \mathbf{y}_{1T_1}^* \\ \mathbf{y}_{21}^* \\ \vdots \\ \mathbf{y}_{2T_2}^* \\ \vdots \\ \mathbf{y}_{N1}^* \\ \vdots \\ \mathbf{y}_{NT_N}^* \end{bmatrix} = \begin{bmatrix} W_{11} \\ \vdots \\ W_{1T_1} \\ W_{21} \\ \vdots \\ W_{2T_2} \\ \vdots \\ W_{N1} \\ \vdots \\ W_{NT_N} \end{bmatrix} \theta + \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_N \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{11} \\ \vdots \\ \boldsymbol{\varepsilon}_{1T_1} \\ \boldsymbol{\varepsilon}_{21} \\ \vdots \\ \boldsymbol{\varepsilon}_{2T_2} \\ \vdots \\ \boldsymbol{\varepsilon}_{N1} \\ \vdots \\ \boldsymbol{\varepsilon}_{NT_N} \end{bmatrix}, \quad (10)$$

or

$$\mathbf{y}^* = W \cdot \theta + \mathbf{u} + \boldsymbol{\varepsilon}, \quad (11)$$

where  $\mathbf{y}^*$  is  $(N \cdot J \cdot T \times 1)$ ,  $W$  is  $(N \cdot J \cdot T \times K)$ ,  $\theta$  is  $(K \times 1)$ ,  $\mathbf{u}$  is  $(N \cdot J \cdot T \times 1)$  with the same value for the  $i^{\text{th}}$  household over all  $T_i$  periods,  $\boldsymbol{\varepsilon}$  is  $(N \cdot J \cdot T \times 1)$ ,  $T = \sum_{i=1}^N T_i$ , and  $K$  is the total number of demand parameters to be estimated.

The individual equations of the SUR tobit model are comprised of parameters that vary across both commodities, households, and time. Using the media index as a proxy for food safety information, the model is estimated for each of the three commodities of interest using the following specification:

$$\begin{aligned} q_{ijt} = & \sum_{j=1}^3 \gamma_j Price_{ijt} + \sum_{j=1}^3 \beta_j MI_{ijt} + \eta Ed_i * MI_{ijt} + \delta Age_i * MI_{ijt} \\ & + \mu Urban_i * MI_{ijt} + \rho Child_i * MI_{ijt} + \sum_{d=1}^D \tau^d h_i^d, \end{aligned} \quad (12)$$

where  $q_{ijt}$  is the quality-adjusted per capita quantity of commodity  $j$  purchased by household  $i$  in time period  $t$  (can be positive or zero),  $D$  indexes the total number of demographic variables included in the model, and  $h_i^d$  is the  $d$ th demographic characteristic of household  $i$  in time period  $t$ .<sup>6</sup>

The variable *Price* used in the three demand equations is the share-weighted geometric price index for each of the three commodities. The expected impact of *Price* on the probability of purchasing a commodity should be negative. That is, it would be expected that as the price of a good decreases, the probability of a household purchasing it would increase. The expected sign on the prices of the other goods in the model is positive, indicating that the three meat and poultry commodities are substitute goods. The food safety information variable, *MI*, is the commodity- and region-specific media index. The expected effect of an increase in the amount of food safety information available to the public would decrease the probability of purchase for some or possibly all households.

Interaction terms between the food safety variable and select demographic variables are included in the model. The education variable, *Ed*, used in the model is a binary variable equal to one if the head of household has a college or post college education and zero otherwise.<sup>7</sup> *Age* is measured as a binary variable equal to one if the head of household is aged 55 or older and zero otherwise. The effect of children, *Child*, is measured using a binary variable equal to one if children under the age of 18 are present in the household and zero otherwise. The final demographic variable used in the interaction terms with food safety information, *Urban*, is a

---

<sup>6</sup> The demographic variables such as age, education, and race do not vary over time. However, the notation also includes the binary variables for annual and monthly seasonal effects, which do vary over time.

<sup>7</sup> Demographic information is provided for both the male and female in married households, but no designation is made for the primary person responsible for purchase decisions. Therefore, it was arbitrarily decided that the demographic information for the female head of household would be used in model estimation.

binary variable indicating the location of the household in an urban area. *Urban* equals one if the household resides in an urban area and equals zero otherwise. The demographic variables for children and head of household aged 55 and older are used in the food safety interactions because these two groups of people are potentially the most susceptible to serious illness from foodborne pathogens. The education dummy variable is included to reflect possible differences in the gathering and processing of media information between households with and without college degrees. Finally, the urban location variable is interacted with food safety information to reflect possible differences information dissemination between urban and rural areas. For example, the limited availability of cable television or high speed internet connections in rural areas may impact the type and quantity of information that rural households will receive. There are no a priori expectations of the effect of the interaction terms on the probability of purchasing the three commodities. In addition to the interaction terms, the select household demographic variables of *Ed*, *Age*, *Child*, and *Urban* also enter the model separately to account for the average effects of these characteristics.

Other variables included in the binary choice models are household specific. They include variables for household income, *Income*, and a quadratic household income term,  $Income^2$ . The expected effect of income on the probability of purchasing beef, pork, or poultry is positive, while the expected sign for the squared term is negative. This reflects a positive, but declining effect of income on the probability of meat and poultry purchases.<sup>8</sup> The size of the household is also included in the regression (*Hsize*) to account for possible differences in purchase patterns for large versus small families. Seasonal effects in the purchase patterns of

---

<sup>8</sup> The household income data were scaled by dividing each observation by 10,000. Therefore, the coefficients for the income variables can be interpreted as the change in the dependent variable caused by a change in total household income of \$10,000.

households are accounted for using monthly dummy variables ( $M1-M12$ ) with the parameter for December ( $M12$ ) omitted from the regression. Annual effects in demand are also considered using year dummy variables ( $Y1-Y8$ ) with the variable for 2003 ( $Y6$ ) omitted from the regression. The expected signs for these variables are not known a priori, but are expected to vary by commodity. The geographic location of the household is included as binary variables for the central, western, and northeastern regions ( $Central, West, Northeast$ ) with the variable for the southern region dropped from the regression. The race of the head of household is categorized into Caucasian, Hispanic, black, Asian, and Other race. The variables  $Hispanic, Black, Asian,$  and  $Other$  are included in the model and the variable  $Caucasian$  is omitted. The expected signs of the geographic location and race variables are not known a priori.

### Estimation Methodology

The SUR tobit model is a generalization of the single equation tobit model. The primary estimation difficulty with SUR tobit is that as the number of equations (commodities) increases, the model becomes more difficult to estimate. This is due to the increase in the number of possible censored commodities. For example, if there are  $p$  commodities (equations), then there would be  $2^p$  possible combinations of censored commodities. Using Huang's (2001) notation, the  $2^p$  possible combinations may be represented by the following  $2^p \times 1$  vector:

$$S = \left[ S_1 = (0, \dots, 0)', \dots, S_h = \left( \underbrace{0, \dots, 0}_r, \underbrace{+, \dots, +}_{p-r} \right)', \dots, S_{2^p} = (+, \dots, +)' \right], \quad (13)$$

where  $S_k$  is  $(p \times 1)$ ,  $k = 1, 2, \dots, 2^p$ ,  $r$  is the number of censored commodities, '+' indicates a positive purchase level for the commodity, and '0' implies a censored observation for the

commodity in the random effects SUR tobit model. The likelihood function for the  $i^{\text{th}}$  household in the  $S_h$  case is given by:

$$L_i^{S_h} (y_i | W, \Sigma) = \int_{-\infty}^{-W_i \theta_1} \dots \int_{-\infty}^{-W_i \theta_r} \left\{ (2\pi)^{-p/2} |\Sigma^{-1}|^{1/2} \exp -\frac{1}{2} (y_i^* - W_i \theta)' \Sigma^{-1} (y_i^* - W_i \theta) \right\} . \quad (14)$$

It is clear that as the number of censored commodities approaches  $2^p$ , the dimension of integration increases. In systems with large numbers of equations, this likelihood function quickly becomes intractable.<sup>9</sup>

Given the complexities of estimation when censoring is present in a SUR model, it may be advantageous to use a methodology that augments or ‘fills in’ the latent dependent variables during estimation, thereby avoiding the need to compute integrated probabilities. This would simplify estimation to that of a standard non-censored SUR model. This study will employ a Bayesian analysis that allows for the use of a data augmentation methodology nested within a Gibbs sampler routine for posterior simulation. The Gibbs sampler was first introduced by Geman and Geman (1984) and a general explanation of the technique is found in Casella and George (1992). It is a Markov Chain Monte Carlo (MCMC) approach that generates random draws of variables from complex multivariate distributions by sampling sequentially from the full set of conditional distributions. The Gibbs sampler was shown by Percy (1992) to be suitable for estimation of the SUR model in a Bayesian analysis. Chib (1992) incorporated the idea of data augmentation into a Gibbs sampler for estimation of a single equation tobit model and the approach was extended to the SUR tobit model by Huang (2001). See Appendix 1 for details of the Bayesian estimation of the random effects SUR tobit model employed in this study.

---

<sup>9</sup> Several alternative methodologies for estimating systems of censored demand equations have been put forth in the literature (e.g. Dong, Gould, and Kaiser (2004); Perali and Chavas (2000); Golan, Perloff, and Shen (2001)). The techniques used in these studies vary widely, suggesting that a general consensus on estimation methodology does not exist.

## Results

Due to the large size of the dataset and the amount of time needed to run these models, a subsample of the data was used for estimation. A random sample of 3,000 households was selected from the original dataset. All the observations from the panel were used for each of the 3,000 households. This resulted in 119,280 observations that were used for estimation. Summary statistics are presented in table 3 for both the full dataset and the random sample.

Bayesian coefficients are typically the mean of the posterior samples. Drawing from the Bernstein-von Mises theorem, the posterior analysis presented here is given a classical statistical interpretation.<sup>10</sup> The classical perspective allows for discussion of the ‘statistical significance’ of the coefficients using confidence intervals. Summarizing the upper and lower 2.5% tails of the posterior distributions gives 95% confidence intervals for each parameter. Coefficients with confidence intervals that do not contain zero are referred to as statistically significantly different from zero. Results of the random effects SUR tobit model are presented in table 4. The means, standard deviations, and 95% confidence intervals are calculated using 1,015 posterior realizations.

Rather than interpret the signs and statistical significance of the parameter estimates, household-level elasticities for prices, income, and food safety for the various demographic subgroups are discussed. Elasticities are useful for several reasons. First, although some of the food safety media index interaction terms with the demographic subgroups are statistically significant, the total effect for these subgroups (the average media effect plus the interaction coefficient) may or may not also be statistically significantly different from zero. Calculation of

---

<sup>10</sup> The Bernstein-von Mises theorem states that as the sample size increases, the posterior distribution becomes normal and the variance of the posterior becomes the same as the sampling variance of the maximum likelihood estimator, implying that the mean of the posterior distribution (the Bayesian coefficients) is asymptotically equivalent to the maximum likelihood estimate (Train, pp.291-293, 2003).



the total food safety elasticity for each realization of the parameter vector will give both an average elasticity as well as the standard deviation. This provides more information about the statistical significance of the total effect for food safety. Second, elasticities provide estimates of purchase response that is unitless. This allows for a comparison of the effects of prices and income relative to food safety information.

The elasticities are calculated using the marginal effects rather than the parameter estimates. The estimates of the unknown parameters are defined as follows:

$$\boldsymbol{\theta} = \frac{\partial E(y_i^*)}{\partial W_i}, \quad (15)$$

where  $i$  denotes an individual household. The parameter estimates reflect the changes in the mean of the latent dependent variable for a change in an independent variable. The marginal effects are:

$$\mathbf{m} = \frac{\partial E(y_i)}{\partial W_i}, \quad (16)$$

and reflect the changes in the unconditional expected values of the observed dependent variable for a change in the independent variables. The use of the marginal effects allows the elasticities to be calculated using the full sample means for the regressors ( $W_i$ ) and the mean of the dependent variable for positive purchases only ( $y_i$ ). The marginal effects for the  $i$ th household and the  $j$ th equation of the random effects SUR tobit model are calculated as:

$$m_{ij} = \Phi \left( \frac{(\alpha_j + \mathbf{x}_{ij}\boldsymbol{\beta}_j + \mathbf{c}_i\boldsymbol{\gamma}_j)}{(\Sigma_{jj} + V_{jj})^{\frac{1}{2}}} \right) \theta_j, \quad (17)$$

where  $V_{jj}$  is the  $j$ th diagonal element of the household-specific error variance matrix. For the random effects model, the marginal effects of the  $j$ th equation,  $\mathbf{m}_j$ , are calculated as the average over all the posterior realizations.

The own-price elasticity of the  $j$ th commodity is calculated as follows:

$$E_j^{price} = m_j^{price} * \frac{\bar{p}_j}{\bar{y}_j}, \quad (18)$$

where  $m_j^{price}$  is the own-price marginal effect for the  $j$ th commodity,  $\bar{p}_j$  is the mean price calculated over the full sample of households and  $\bar{y}_j$  is the mean quantity calculated using only the positive purchases of the  $j$ th commodity. The cross-price elasticity is calculated as follows:

$$E_{jl}^{price} = m_{jl}^{price} * \frac{\bar{p}_l}{\bar{y}_j} \text{ for } j \neq l, \quad (19)$$

where  $m_{jl}^{price}$  is the cross-price marginal effect for the  $j$ th commodity. The income elasticity is calculated as follows:

$$E_j^{inc} = \left( m_j^{inc} + 2 * m_j^{inc^2} * \overline{inc} \right) * \frac{\overline{inc}}{\bar{y}_j}, \quad (20)$$

where  $m_j^{inc}$  is the income marginal effect for the  $j$ th commodity,  $m_j^{inc^2}$  is the income squared marginal effect for the  $j$ th commodity, and  $\overline{inc}$  is the mean household income calculated over the full sample of households.

The elasticity of quantity purchased with respect to the media index is similarly calculated for each commodity and demographic subgroup. The formula for the food safety elasticity with respect to education is as follows:

$$E_j^{MI*Ed} = \left( m_j^{MI} + m_j^{MI*Ed} \right) * \frac{\overline{MI}_j}{\bar{y}_j^{Ed}}, \quad (21)$$

where  $m_j^{MI}$  is the coefficient for food safety of the  $j$ th commodity,  $m_j^{MI*Ed}$  is the marginal effect of the interaction term between the  $j$ th commodity media index and the dummy variable for a college educated head of household, and  $\overline{MI}_j$  is the mean value of the media index variable for the  $j$ th commodity calculated using only the college educated head of household subgroup. The food safety elasticities for the other demographic subgroups (age 55 and older head of household, children present in the household, and household located in urban area) are similarly calculated.

The price and food safety elasticities are presented in table 5 for the random effects model. All of the own-price elasticities are statistically different from zero using a 95% confidence interval. The own-price elasticities are greater than one for beef and pork, but relatively inelastic for pork. The beef price elasticity indicates that a 10% increase in the price of beef would cause a 13.0% decline in per capita beef purchases. The effect from a 10% increase in the price of pork is estimated to be a 6.9% decline in purchases. The price effect for poultry is very comparable to that of beef price with an estimated decrease of 15.1% from a 10% increase in price. All but one of the cross-price elasticities for beef, pork, and poultry are statistically significantly different from zero and have negative signs. The cross-price elasticity of pork price on poultry purchases is not statistically significant. The cross-price elasticities are small in magnitude as compared to the own-price elasticities suggesting that a change in the price of another good in the system has very limited impact on the quantity purchased of the other goods.

The elasticities with respect to income are statistically significant for all three commodities. For beef, a 10% increase in household income increases the pounds per capita purchased by 1.6%. The effects for pork and poultry are increases in per capita purchases of

0.8% and 1.7%, respectively. These effects are similar in magnitude as compared to the cross-price effects, but are much smaller than the own-price effects.

The food safety elasticities for households located in urban areas are statistically significantly different from zero for every commodity media index. The effect of a 10% increase in the poultry index is estimated to be a decrease in purchases of 0.4% for these households. However, an increase in the beef and pork media indices is estimated to cause a 0.1% increase in the amount of beef and pork urban household purchase. All the remaining food safety elasticities are not significantly different from zero. The food safety effects that are statistically significant are relatively small in magnitude and do not appear to be as economically significant as the price and income elasticities.

The price and food safety elasticities estimated in this study are comparable to elasticity estimates given in other studies. A literature search conducted by the U.S. Environmental Protection Agency (pg. 3-41, 2002) indicated the following ranges of own-price elasticities for meat and poultry: -2.590 to -0.150 for beef; -1.234 to -0.070 for pork; -1.250 to -0.104 for broilers; and -0.680 to -0.372 for turkeys. The own price elasticity estimates from the random effects model for beef and pork fall within these ranges.<sup>11</sup> The relatively high magnitude of the poultry price effect is similar to the results found by Piggott and Marsh (2004). They found that pre-committed quantities of beef and pork were higher than for poultry, suggesting that poultry purchases may be more sensitive to changes in price and income than beef and pork purchases. The food safety elasticities estimated in the Piggott and Marsh study are -0.0144 for beef, -0.0131 for pork, and -0.0250 for poultry. These elasticities measure the total effect of food safety information on the representative consumer. The magnitudes of their elasticities are very

---

<sup>11</sup> The own-price elasticities for poultry fall outside the ranges for both broilers and turkey. However, the use of a poultry aggregate, which includes both chicken and turkey products, in this study may explain this difference in estimated elasticities.

comparable to the food safety elasticities found in this study for each of the four demographic groups of households.

## **Conclusion**

The elasticities calculated from the results of the random effects SUR tobit model indicate that food safety information does not have a statistically significant effect for the vast majority of the households considered in the model. The only statistically significant effects were for households in urban areas. For the few food safety elasticities in the random effects SUR tobit model there are statistically significant, their small magnitude relative prices and income indicates that they are not necessarily important economically.

The results of this study are similar to previous research. Piggott and Marsh (2004) found statistically significant food safety effects, but they were small in magnitude and short-lived. However, their study used aggregate disappearance data to measure consumption. These data include consumption of meat and poultry both at home and away from home. The data employed in this study only account for food purchased for consumption at home. Therefore, differences in the statistical significance of food safety information between the Piggott and Marsh study and the results presented here may be due in part to differences in the consumption measure employed. While the Schlenker and Villas-Boas (2006) study found statistically significant effects at the grocery store level, it did not find these same effects at the household level for meat purchases. One possible reason that results at the store and household levels differ is that the aggregation of product groups for household purchases may mask the product substitution that is noticeable at the store level.

The elasticities calculated from the results of the random effects SUR tobit model indicate that food safety information does not have a statistically significant effect on purchases

of meat and for the vast majority of the households considered in the model. However, households located in urban areas have a statistically significant response which is negative for poultry and positive for beef and pork purchases. A negative effect from food safety information is an intuitive result. It implies that people will decrease their purchases of poultry, probably in favor of other foods. However, a slightly positive response to beef and pork food safety information is not necessarily an implausible response. Many food safety recalls are product specific, impacting only ground beef, for example. Consumers may still continue to buy other beef products, like roasts or steaks, but avoid purchasing ground beef. As a result, their overall purchases of beef may not change or could even increase slightly, while still responding rationally to the food safety information with regard to ground beef. These results suggest that further investigation of heterogeneous household effects using different aggregation levels of meat and poultry products is warranted.

One aspect of consumer behavior that was not explicitly accounted for in this study is the effect of decisions made in previous time periods on the probability of purchase in the current period. The effects from these past decisions can be captured using state dependence variables which can capture both inventory and purchase habit effects. By explaining the variability due to state dependence, second-order effects from food safety information may be more accurately identified.

Future research will also focus on different specifications of the media index. For example, the specification of a 30-day rolling average using a two-week memory has an intuitive appeal given the frequency with which household make meat and poultry purchases. However, it is possible that a longer lag length or a distributed lag structure would be a better fit for the data. The most appropriate specification of the lag structure of the media index is an empirical

question that remains to be answered. Other specifications could focus on the criteria applied to article searches. Currently, any article pertaining to meat or poultry and food safety that is found in the regional newspapers is used, including articles focused on international events. If consumer purchase decisions are not impacted by international events, then the current media index specification may be inappropriate. An alternative to this specification would be to use only those articles that focus on domestic food safety events or issues. While there are an endless number of specifications for the media index, each specification that is analyzed provides researchers with more information on how to model consumer behavior and food safety information.

**Table 1 Household Panel Demographic Variables**

<b>Demographic Variable</b>	<b>Frequency</b>	<b>Percent of Sample <sup>a</sup></b>
<b>Household Size</b>		
Single member	1,820	23.33
Two members	2,913	37.48
Three members	1,222	15.76
Four members	1,087	14.05
Five members	479	6.19
Six members	160	2.06
Seven members	57	0.74
Eight members	18	0.23
Nine or more members	13	0.17
<b>Household Income</b>		
Under \$5000	46	0.59
\$5000-\$7999	73	0.94
\$8000-\$9999	72	0.93
\$10,000-\$11,999	107	1.37
\$12,000-\$14,999	198	2.54
\$15,000-\$19,999	388	4.99
\$20,000-\$24,999	559	7.19
\$25,000-\$29,999	496	6.40
\$30,000-\$34,999	581	7.48
\$35,000-\$39,999	541	6.97
\$40,000-\$44,999	584	7.55
\$45,000-\$49,999	528	6.81
\$50,000-\$59,999	901	11.63
\$60,000-\$69,999	767	9.89
\$70,000-\$99,999	1,223	15.72
\$100,000 & Over	705	9.02
<b>Age of Male Head <sup>b</sup></b>		
Under 25 Years	23	0.30
25-29 Years	160	2.09
30-34 Years	431	5.58
35-39 Years	608	7.85
40-44 Years	719	9.29
45-49 Years	791	10.22
50-54 Years	760	9.82
55-64 Years	1,210	15.56
65+ Years	1,079	13.82
No Male Head	1,987	25.48
<b>Age of Female Head <sup>b</sup></b>		
Under 25 Years	52	0.69
25-29 Years	250	3.26
30-34 Years	549	7.11
35-39 Years	730	9.44
40-44 Years	889	11.49
45-49 Years	966	12.47
50-54 Years	951	12.24
55-64 Years	1,467	18.80
65+ Years	1,158	14.82
No Female Head	755	9.70

<sup>a</sup> Summary statistics calculated as average across the eight sample years.

<sup>b</sup> Married households have information on both the male and female head of household.



**Table 1 Household Panel Demographic Variables, cont.**

<b>Demographic Variable</b>	<b>Frequency</b>	<b>Percent of Sample <sup>a</sup></b>
<b>Age and Presence of Children</b>		
Under 6 only	330	4.29
6-12 only	549	7.08
13-17 only	628	8.13
Under 6 & 6-12	302	3.90
Under 6 & 13-17	48	0.61
6-12 & 13-17	372	4.80
Under 6 & 6-12 & 13-17	67	0.87
No Children Under 18	5,472	70.33
<b>Male Head Employment <sup>b</sup></b>		
Under 30 hours	235	3.02
30-34 hours	140	1.80
35+ hours	3,937	50.89
Not Employed for Pay	1,468	18.81
No Male Head	1,987	25.48
<b>Female Head Employment <sup>b</sup></b>		
Under 30 hours	885	11.41
30-34 hours	378	4.88
35+ hours	3,203	41.34
Not Employed for Pay	2,547	32.68
No Female Head	755	9.70
<b>Male Head Education <sup>b</sup></b>		
Grade School	76	0.97
Some High School	291	3.74
Graduated High School	1,315	16.93
Some College	1,767	22.79
Graduated College	1,548	19.97
Post College Grad	783	10.12
No Male Head	1,987	25.48
<b>Female Head Education <sup>b</sup></b>		
Grade School	38	0.48
Some High School	206	2.65
Graduated High School	1,765	22.70
Some College	2,376	30.61
Graduated College	1,892	24.37
Post College Grad	737	9.50
No Female Head	755	9.70
<b>Region</b>		
East	1,658	21.32
Central	1,582	20.53
South	2,840	36.45
West	1,687	21.70
<b>Marital Status</b>		
Married	4,755	61.37
Widowed	618	7.90
Divorced/Separated	1,142	14.64
Single	1,253	16.09

<sup>a</sup> Summary statistics calculated as average across the eight sample years.

<sup>b</sup> Married households have information on both the male and female head of household.

a

**Table 2 Summary Statistics of Quality-Adjusted Monthly Purchases and Price Indices**

	<b>Average</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Std. Dev.</b>
<b>Beef</b>				
Per Capita Quantity (lbs)	4.901	0	1,452.640	8.584
Geometric price index	3.046	0.170	8.006	0.493
<b>Pork</b>				
Quantity (lbs)	2.129	0	408.725	5.159
Geometric price index	2.480	0.055	10.795	0.476
<b>Poultry</b>				
Quantity (lbs)	3.101	0	1,911.060	6.468
Geometric price index	1.822	0.150	6.045	0.245

<sup>a</sup>Summary statistics based on 745,632 monthly observations.

**Table 3 Summary Statistics of Demand Model Variables**

	Full Sample				Random Sample			
	Average	Minimum	Maximum	Std. Dev.	Average	Minimum	Maximum	Std. Dev.
<b>Beef Price</b>	3.209	0.577	12.638	0.562	3.196	1.227	12.638	0.551
<b>Pork Price</b>	2.534	0.627	12.219	0.509	2.527	0.644	11.453	0.513
<b>Poultry Price</b>	1.924	0.700	8.195	0.248	1.918	0.880	7.082	0.248
<b>Beef MI</b>	7.633	0.786	77.645	6.428	7.650	0.786	77.645	6.446
<b>Pork MI</b>	2.547	0.000	16.567	1.988	2.558	0.000	16.567	2.010
<b>Poultry MI</b>	11.378	2.000	38.310	6.054	11.336	2.000	38.310	6.021
<b>Ed</b>	0.393	0	1	0.488	0.376	0	1	0.484
<b>Age</b>	0.372	0	1	0.483	0.376	0	1	0.484
<b>Urban</b>	0.875	0	1	0.330	0.873	0	1	0.333
<b>Child</b>	0.296	0	1	0.456	0.288	0	1	0.453
<b>Income</b>	5.383	0.250	12.500	3.151	5.281	0.250	12.500	3.137
<b>Income<sup>2</sup></b>	38.910	0.062	156.250	43.477	37.729	0.062	156.250	43.064
<b>Y1</b>	0.120	0	1	0.325	0.120	0	1	0.325
<b>Y2</b>	0.112	0	1	0.316	0.114	0	1	0.318
<b>Y3</b>	0.118	0	1	0.322	0.118	0	1	0.323
<b>Y4</b>	0.127	0	1	0.333	0.130	0	1	0.337
<b>Y5</b>	0.133	0	1	0.340	0.131	0	1	0.338
<b>Y6</b>	0.136	0	1	0.342	0.134	0	1	0.341
<b>Y7</b>	0.129	0	1	0.336	0.130	0	1	0.336
<b>Y8</b>	0.125	0	1	0.330	0.122	0	1	0.328
<b>M1</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M2</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M3</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M4</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M5</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M6</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M7</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M8</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M9</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M10</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M11</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>M12</b>	0.083	0	1	0.276	0.083	0	1	0.276
<b>South</b>	0.366	0	1	0.482	0.362	0	1	0.481
<b>Central</b>	0.204	0	1	0.403	0.216	0	1	0.412
<b>West</b>	0.217	0	1	0.412	0.216	0	1	0.412
<b>Northeast</b>	0.213	0	1	0.410	0.205	0	1	0.404
<b>Caucasian</b>	0.766	0	1	0.423	0.758	0	1	0.429
<b>Hispanic</b>	0.076	0	1	0.264	0.075	0	1	0.264
<b>Black</b>	0.121	0	1	0.326	0.123	0	1	0.328
<b>Asian</b>	0.022	0	1	0.146	0.026	0	1	0.159
<b>Other</b>	0.016	0	1	0.126	0.018	0	1	0.134

Note: The number of observations in the full sample is 745,632 and the number of observations in the random sample of 3,000 households is 119,280.

**Table 4 Bayesian Estimated Coefficients of the Random Effects SUR Tobit Model**

	Beef Model				Pork Model				Poultry Model			
	Coefficient	Standard Deviation	95% Confidence Interval		Coefficient	Standard Deviation	95% Confidence Interval		Coefficient	Standard Deviation	95% Confidence Interval	
<b>Beef Price</b>	-7.899	0.113	-8.110	-7.676	-0.452	0.106	-0.675	-0.240	-0.600	0.097	-0.795	-0.412
<b>Pork Price</b>	-0.493	0.117	-0.731	-0.268	-5.616	0.087	-5.788	-5.450	-0.175	0.093	-0.357	0.004
<b>Poultry Price</b>	-1.044	0.231	-1.503	-0.603	-0.692	0.197	-1.069	-0.322	-13.093	0.169	-13.444	-12.761
<b>Beef MI</b>	0.027	0.022	-0.018	0.071	0.000	0.007	-0.013	0.013	0.007	0.006	-0.005	0.020
<b>Pork MI</b>	0.005	0.028	-0.049	0.061	0.080	0.056	-0.030	0.193	0.056	0.021	0.014	0.096
<b>Poultry MI</b>	0.003	0.010	-0.018	0.022	-0.011	0.008	-0.029	0.005	-0.020	0.024	-0.068	0.026
<b>Ed*MI<sub>beef</sub></b>	-0.045	0.012	-0.069	-0.021	--	--	--	--	--	--	--	--
<b>Age*MI<sub>beef</sub></b>	-0.024	0.013	-0.050	0.001	--	--	--	--	--	--	--	--
<b>Child*MI<sub>beef</sub></b>	-0.030	0.015	-0.060	-0.001	--	--	--	--	--	--	--	--
<b>Urban*MI<sub>beef</sub></b>	0.002	0.018	-0.037	0.038	--	--	--	--	--	--	--	--
<b>Ed*MI<sub>pork</sub></b>	--	--	--	--	-0.094	0.037	-0.167	-0.023	--	--	--	--
<b>Age*MI<sub>pork</sub></b>	--	--	--	--	-0.106	0.037	-0.179	-0.033	--	--	--	--
<b>Child*MI<sub>pork</sub></b>	--	--	--	--	-0.098	0.040	-0.179	-0.015	--	--	--	--
<b>Urban*MI<sub>pork</sub></b>	--	--	--	--	0.014	0.049	-0.088	0.105	--	--	--	--
<b>Ed*MI<sub>poultry</sub></b>	--	--	--	--	--	--	--	--	0.008	0.013	-0.017	0.032
<b>Age*MI<sub>poultry</sub></b>	--	--	--	--	--	--	--	--	0.038	0.013	0.011	0.065
<b>Child*MI<sub>poultry</sub></b>	--	--	--	--	--	--	--	--	0.005	0.015	-0.024	0.033
<b>Urban*MI<sub>poultry</sub></b>	--	--	--	--	--	--	--	--	-0.034	0.022	-0.076	0.010
<b>Ed</b>	-0.850	0.278	-1.384	-0.293	-0.504	0.212	-0.932	-0.083	0.021	0.231	-0.439	0.454
<b>Age</b>	0.895	0.217	0.465	1.329	1.112	0.194	0.734	1.479	-0.133	0.221	-0.566	0.303
<b>Child</b>	-2.699	0.208	-3.127	-2.299	-1.413	0.194	-1.786	-1.034	-2.053	0.227	-2.520	-1.617
<b>Urban</b>	0.188	0.325	-0.462	0.833	-0.204	0.276	-0.751	0.315	1.215	0.339	0.536	1.877
<b>Income</b>	0.796	0.098	0.610	0.992	0.502	0.079	0.352	0.658	0.493	0.077	0.335	0.643
<b>Income<sup>2</sup></b>	-0.023	0.006	-0.037	-0.011	-0.019	0.005	-0.031	-0.009	0.000	0.005	-0.010	0.010

Note: The estimated coefficients are means calculated from 1,015 posterior realizations. The 95% confidence intervals are calculated using the upper and lower 2.5% of the posterior distribution.

**Table 4 Bayesian Estimated Coefficients of the Random Effects SUR Tobit Model, cont.**

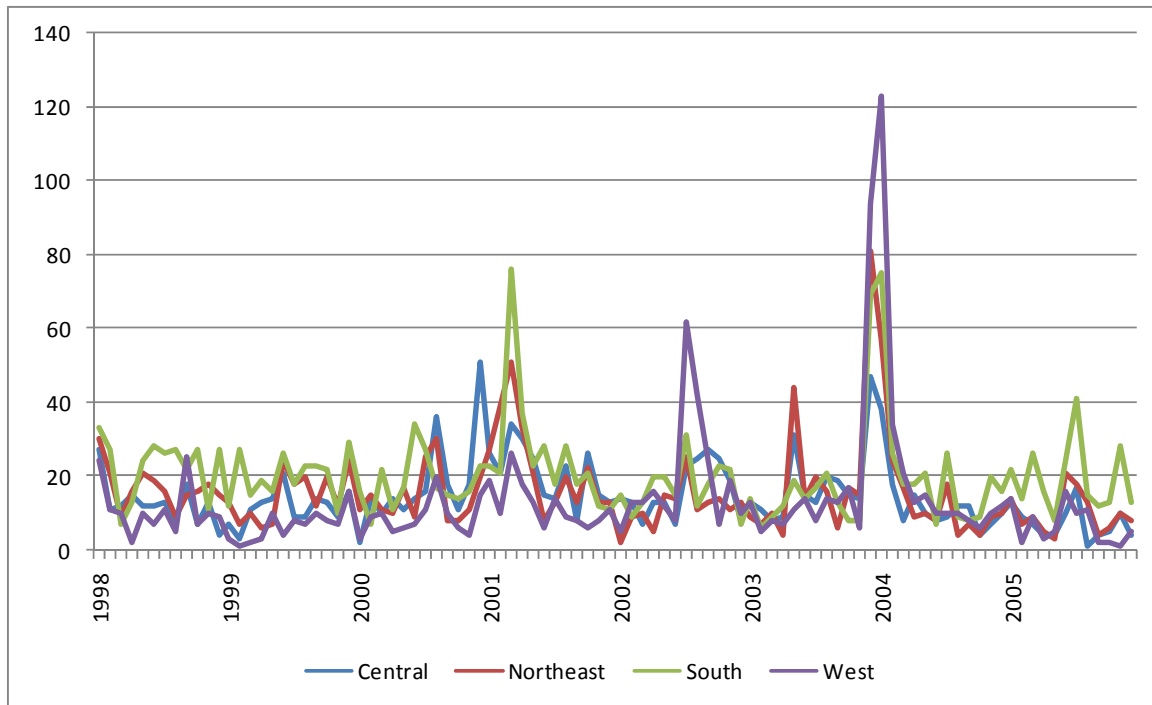
	Beef Model				Pork Model				Poultry Model			
	Coefficient	Standard Deviation	95% Confidence Interval		Coefficient	Standard Deviation	95% Confidence Interval		Coefficient	Standard Deviation	95% Confidence Interval	
<b>Y1</b>	0.265	0.223	-0.199	0.675	6.617	0.181	6.252	6.969	4.383	0.176	4.030	4.727
<b>Y2</b>	-1.453	0.179	-1.791	-1.087	1.170	0.150	0.870	1.468	0.442	0.156	0.134	0.748
<b>Y3</b>	-0.081	0.170	-0.404	0.260	0.868	0.142	0.588	1.133	1.207	0.137	0.930	1.480
<b>Y4</b>	-0.331	0.166	-0.661	0.003	0.897	0.137	0.640	1.188	0.349	0.130	0.084	0.597
<b>Y5</b>	-0.713	0.146	-1.010	-0.429	-0.200	0.122	-0.435	0.045	-0.452	0.124	-0.701	-0.225
<b>Y7</b>	1.109	0.159	0.808	1.414	0.487	0.129	0.217	0.731	0.724	0.123	0.461	0.952
<b>Y8</b>	0.043	0.173	-0.299	0.392	0.513	0.140	0.222	0.785	1.352	0.131	1.090	1.600
<b>M1</b>	-0.148	0.196	-0.533	0.240	-2.269	0.167	-2.587	-1.938	0.967	0.158	0.670	1.277
<b>M2</b>	-0.407	0.190	-0.778	-0.048	-2.635	0.155	-2.948	-2.349	0.823	0.148	0.528	1.118
<b>M3</b>	0.328	0.187	-0.043	0.679	-1.812	0.147	-2.105	-1.527	0.945	0.150	0.661	1.243
<b>M4</b>	-0.268	0.186	-0.633	0.104	-1.326	0.147	-1.617	-1.035	0.589	0.148	0.296	0.875
<b>M5</b>	1.335	0.191	0.975	1.696	-2.466	0.148	-2.743	-2.174	1.298	0.153	0.998	1.591
<b>M6</b>	0.479	0.192	0.096	0.866	-2.863	0.154	-3.179	-2.560	0.841	0.151	0.530	1.130
<b>M7</b>	0.457	0.191	0.080	0.830	-2.697	0.149	-2.996	-2.407	0.909	0.158	0.594	1.211
<b>M8</b>	0.525	0.184	0.173	0.888	-2.703	0.151	-2.982	-2.401	1.291	0.151	1.009	1.594
<b>M9</b>	0.141	0.188	-0.234	0.502	-2.603	0.148	-2.893	-2.326	0.988	0.149	0.698	1.261
<b>M10</b>	0.143	0.187	-0.243	0.496	-2.511	0.148	-2.798	-2.230	0.687	0.146	0.402	0.985
<b>M11</b>	-1.335	0.190	-1.685	-0.965	-1.974	0.152	-2.266	-1.687	1.529	0.145	1.261	1.809
<b>Central</b>	0.084	0.440	-0.776	0.966	0.184	0.350	-0.481	0.880	-1.070	0.336	-1.739	-0.444
<b>West</b>	1.622	0.435	0.804	2.539	-0.535	0.329	-1.158	0.102	1.992	0.329	1.394	2.678
<b>Northeast</b>	1.391	0.425	0.556	2.226	0.296	0.303	-0.330	0.882	1.284	0.293	0.704	1.879
<b>Hispanic</b>	0.703	0.422	-0.110	1.557	0.086	0.335	-0.548	0.733	0.596	0.311	0.033	1.216
<b>Black</b>	-2.306	0.455	-3.187	-1.344	0.502	0.347	-0.209	1.165	1.725	0.329	1.086	2.381
<b>Asian</b>	-2.180	0.667	-3.476	-0.857	0.326	0.520	-0.783	1.297	0.428	0.496	-0.579	1.392
<b>Other</b>	-2.502	0.489	-3.478	-1.558	0.749	0.400	-0.025	1.538	0.021	0.380	-0.718	0.786
<b>Constant</b>	22.693	0.744	21.164	24.127	10.493	0.630	9.274	11.694	17.886	0.657	16.633	19.156
<b>Sigma <math>\sigma</math></b>	12.363	0.034	12.296	12.433	8.695	0.032	8.632	8.756	9.480	0.030	9.421	9.538
<b>Sigma <math>\mu</math></b>	9.123	0.138	8.858	9.387	6.655	0.114	6.446	6.874	6.414	0.096	6.225	6.602

Note: The estimated coefficients are means calculated from 1,015 posterior realizations. The 95% confidence intervals are calculated using the upper and lower 2.5% tails of the posterior distribution.

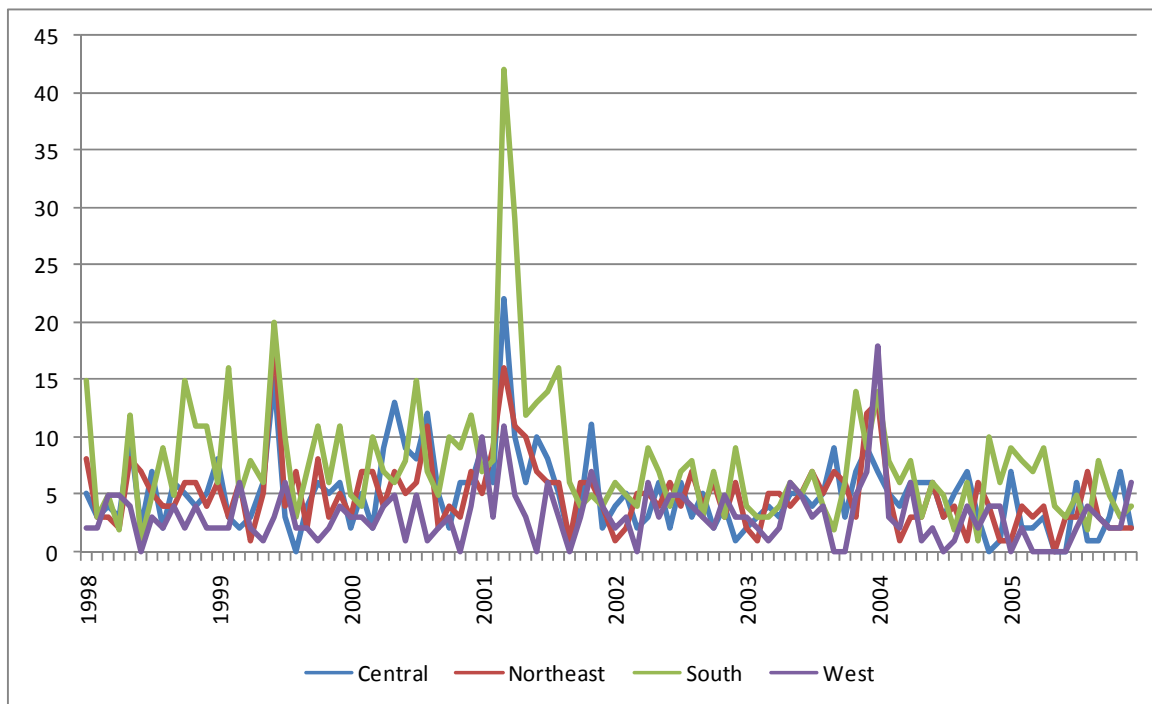
**Table 5 Price and Food Safety Elasticities of Random Effects SUR Tobit Models**

		Elasticity	Standard Deviation	95% Confidence Interval	
<b><u>Own-Price</u></b>					
	<b>Beef</b>	-1.296	0.023	-1.339	-1.251
	<b>Pork</b>	-0.688	0.014	-0.714	-0.662
	<b>Poultry</b>	-1.508	0.024	-1.554	-1.461
<b><u>Cross-Price</u></b>					
<b>Beef</b>	<b>Pork</b>	-0.066	0.015	-0.097	-0.036
	<b>Poultry</b>	-0.103	0.022	-0.148	-0.059
<b>Pork</b>	<b>Beef</b>	-0.068	0.016	-0.099	-0.037
	<b>Poultry</b>	-0.064	0.018	-0.097	-0.029
<b>Poultry</b>	<b>Beef</b>	-0.116	0.018	-0.151	-0.076
	<b>Pork</b>	-0.028	0.014	-0.056	0.000
<b><u>Income</u></b>					
	<b>Beef</b>	0.157	0.011	0.135	0.180
	<b>Pork</b>	0.078	0.009	0.060	0.095
	<b>Poultry</b>	0.165	0.011	0.143	0.186
<b><u>Food Safety</u></b>					
<b>College Education</b>	<b>Beef</b>	-0.008	0.009	-0.028	0.009
	<b>Pork</b>	-0.002	0.008	-0.017	0.013
	<b>Poultry</b>	-0.009	0.018	-0.046	0.024
<b>Age 55 &amp; Older</b>	<b>Beef</b>	0.001	0.007	-0.012	0.015
	<b>Pork</b>	-0.003	0.006	-0.014	0.008
	<b>Poultry</b>	0.011	0.014	-0.019	0.037
<b>Children Present</b>	<b>Beef</b>	-0.001	0.013	-0.028	0.025
	<b>Pork</b>	-0.004	0.012	-0.027	0.019
	<b>Poultry</b>	-0.016	0.028	-0.072	0.038
<b>Urban Residence</b>	<b>Beef</b>	0.012	0.006	0.000	0.022
	<b>Pork</b>	0.012	0.004	0.003	0.021
	<b>Poultry</b>	-0.040	0.009	-0.058	-0.022

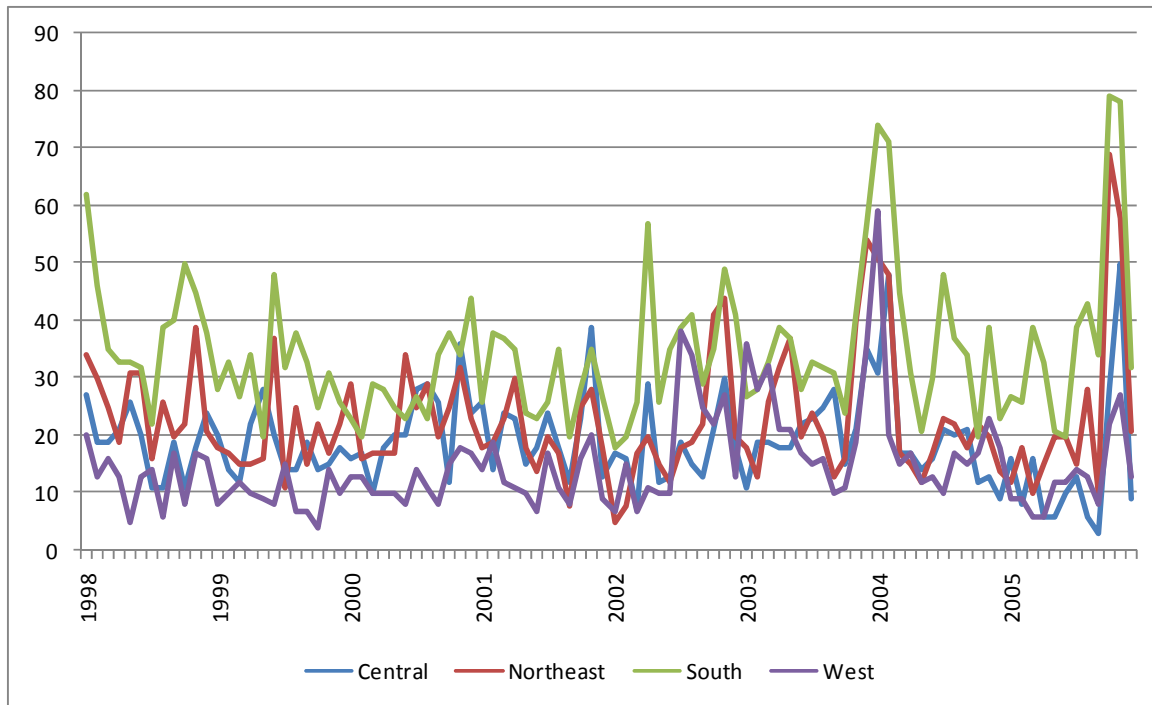
Note: The own- and cross-price elasticities are means calculated from 1,015 posterior realizations. The 95% confidence intervals are calculated using the upper and lower 2.5% tails of the posterior distribution.



**Figure 1 Beef Media Index by Region, 1998 to 2005**



**Figure 2 Pork Media Index by Region, 1998 to 2005**



**Figure 3 Poultry Media Index by Region, 1998 to 2005**



## References

- Burton, M. and T. Young. "The Impact of BSE on the Demand for Beef and Other Meats in Great Britain." *Applied Economics*. 28(1996): 687-693.
- Casella, E. and E.I. George. "Explaining the Gibbs Sampler." *The American Statistician*. 46(1992): 167-174.
- Chib, S. "Bayes Inference in the Tobit Censored Regression Model." *Journal of Econometrics*. 51(1992): 79-99.
- Cox, T. and M. Wohlgenant. "Prices and Quality Effects in Cross-Sectional Demand Analysis." *American Journal of Agricultural Economics*. 68(1986): 908-919.
- Dahlgran, R. and D. Fairchild. "The Demand Impacts of Chicken Contamination Publicity: A Case Study." *Agribusiness*. 18(2002): 459-474.
- Dong, D., B.W. Gould, and H.M. Kaiser. "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(2004): 1094-1107.
- Geman, S. and D. Geman. "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 6(1984): 721-741.
- Golan, A., J.M. Perloff, and E.Z. Shen. "Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand." *Review of Economics and Statistics*. 83(2001): 541-550.
- Huang, H.-C. "Bayesian Analysis of the SUR Tobit Model." *Applied Economics Letters*. 8(2001): 617-622.
- Huber, P. "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions." In *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics*, Vol. 1. Berkeley: University of California Press, 1967.
- Lynch, S.M. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer, 2007.
- Percy, D.F. "Predication for Seemingly Unrelated Regressions." *Journal of the Royal Statistical Society*. 54(1992): 243-252.

- Perali, F. and J.P. Chavas. "Estimation of Censored Demand Equations from Large Cross-Section Data." *American Journal of Agricultural Economics*. 82(2000):1022-1037.
- Piggott, N. and T. Marsh. "Does Food Safety Information Impact U.S. Meat Demand?" *American Journal of Agricultural Economics*. 86(2004): 154-174.
- Schlenker, W. and S.B. Villas-Boas. "Consumer and Market Responses to Mad-Cow Disease." Department of Agricultural and Resource Economics, University of California-Berkeley. CUDARE Working Paper 1023, November 1, 2006.  
[http://repositories.cdlib.org/are\\_ucb/1023/](http://repositories.cdlib.org/are_ucb/1023/)
- Törnqvist, L. "The Bank of Finland's Consumption Price Index." *Bank of Finland Monthly Bulletin*. 10(1936): 1-8.
- Train, K. *Discrete Choice Methods with Simulation*, 40-41. New York: Cambridge University Press, 2003.
- U.S. Environmental Protection Agency – Office of Science and Technology. *Economic Analysis of Proposed Effluent Limitations Guidelines and Standards for the Meat and Poultry Products Industry*. EPA-821-B-01-006. Washington, DC, 2002.
- White, H. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." *Econometrica*, 48(1980):817-838.

# Appendix 1

## Bayesian Estimation: The Random Effects SUR Tobit Model

The estimation of a model in the Bayesian framework requires summarization of a posterior probability distribution. The posterior is derived using Bayes Theorem for probability distributions, which can be stated as:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

where  $\propto$  means “is proportional to.” Given both a likelihood function and prior distributions, a posterior distribution for the unknown model parameters can be derived. The likelihood function is derived from the specification of the model and the prior distributions are determined using any pre-existing knowledge of the model parameters.

The random effects SUR tobit model, stacked over all  $J$  commodities is specified as:

$$\mathbf{y}_{it}^* = \mathbf{W}_{it}\boldsymbol{\theta} + \mathbf{u}_i + \boldsymbol{\varepsilon}_{it} , \quad (\text{A1.1})$$

$$y_{ijt} = \begin{cases} y_{ijt}^* & \text{if } y_{ijt}^* > 0 \\ 0 & \text{if } y_{ijt}^* \leq 0 \end{cases} , \quad (\text{A1.2})$$

where  $\boldsymbol{\varepsilon}_{it} \sim iid N(0, \Sigma)$  and  $\mathbf{u}_i \sim iid N(0, V)$ . The prior distribution of the unknown model parameters,  $\pi(\boldsymbol{\theta})$ , is specified as a multivariate normal distributions. The prior distributions of the unknown parameters,  $\pi(\Sigma)$  and  $\pi(V)$ , are specified as inverse Wishart distributions. The probability distribution of the dependent variable conditional on the model parameters and observed data for household  $i$  in time period  $t$  is:

$$p(\mathbf{y}_{it} | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}_i) = \int_{-\infty}^{-\mathbf{W}'_{it}\boldsymbol{\theta}_1} \dots \int_{-\infty}^{-\mathbf{W}'_{it}\boldsymbol{\theta}_r} f(\mathbf{y}_{it}^* | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}_i) d\mathbf{y}_{it1}^* \dots d\mathbf{y}_{itr}^* , \quad (\text{A1.3})$$

where  $f(\cdot)$  is the normal probability distribution function and  $r$  refers to the number of censored commodities. The likelihood function over all households and time periods is:

$$L(\mathbf{y} | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}) = \prod_{i=1}^N \prod_{t=1}^T p(\mathbf{y}_{it} | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}_i) = p(\mathbf{y} | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}) . \quad (\text{A1.4})$$

Using the likelihood function and prior distributions, the posterior is proportional to the product of the likelihood function and the prior distributions:

$$p(\mathbf{y} | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}) \propto L(\mathbf{y} | \boldsymbol{\theta}, \mathbf{W}, \Sigma, \mathbf{u}) \cdot \pi(\boldsymbol{\theta}) \cdot \pi(\Sigma) \cdot \pi(V) . \quad (\text{A1.5})$$

No analytical form exists for the multivariate posterior distribution given in equation (A1.5), making sampling very difficult. To obtain the conditional posterior distributions needed to employ the Gibbs sampler, the posterior of the unknown model parameters is augmented with

the latent data to get a full posterior. Using properties of probability distributions, the full posterior can be rewritten as follows:

$$\begin{aligned} p(\theta, \Sigma, \mathbf{u}, \mathbf{y}^* | \mathbf{y}, W) &\propto p(\mathbf{y}, \mathbf{y}^* | \theta, W, \Sigma, \mathbf{u}) \cdot \pi(\theta) \cdot \pi(\Sigma) \cdot \pi(V) \\ &\propto p(\mathbf{y} | \mathbf{y}^*, \theta, W, \Sigma, \mathbf{u}) \cdot p(\mathbf{y}^* | \theta, W, \Sigma, \mathbf{u}) \cdot \pi(\theta) \cdot \pi(\Sigma) \cdot \pi(V) . \end{aligned} \quad (\text{A1.6})$$

The conditional posterior distributions are derived using multivariate (univariate) normal-inverse Wishart (gamma) conjugate prior analysis. The Gibbs sampler can now be implemented to sample iteratively from the conditionals in the following order:

$$\begin{aligned} (1) \quad & p(\mathbf{y}^* | \theta, \Sigma, \mathbf{u}, W, \mathbf{y}) \\ (2) \quad & p(V | \mathbf{z}, \theta, \Sigma, \mathbf{u}, W) \\ (3) \quad & p(\mathbf{u} | \mathbf{z}, \theta, \Sigma, V, W) \\ (4) \quad & p(\Sigma | \mathbf{z}, \theta, \mathbf{u}, W) \\ (5) \quad & p(\theta | \mathbf{z}, \Sigma, \mathbf{u}, W) , \end{aligned} \quad (\text{A1.7})$$

where  $\mathbf{z}$  denotes a vector comprised of the observed values of the dependent variable,  $\mathbf{y}$ , and the sampled values of the latent dependent variable,  $\mathbf{y}^*$ .

The truncated normal distribution used in the first step of the Gibbs sampler is conditioned on the household-specific error component  $\mathbf{u}_i$ , which enters the mean of the distribution. Let  $\mathbf{z}_{it} = (\mathbf{y}_{it,r}^*, \mathbf{y}_{it,-r})$  be a vector of dependent variables for the  $i^{\text{th}}$  household with  $r$  denoting elements censored at zero and  $-r$  denoting positive (observed) commodity purchases. The conditional distribution of  $\mathbf{y}_{it,r}^*$  is a truncated normal distribution of the following form:

$$\mathbf{y}_{it,r}^* | \theta, \Sigma, \mathbf{u}_i, W_{it}, \mathbf{y}_{it,-r} \sim TN_{(-\infty, 0]}(\boldsymbol{\mu}_{it,r}, \Sigma_r) , \quad (\text{A1.8})$$

where  $\mathbf{y}_{it,r}^*$  is a dimension  $(r \times 1)$  vector of draws and  $\mathbf{y}_{it,-r}$  is a  $((J-r) \times 1)$  dimension vector of positive purchases. For the  $i^{\text{th}}$  household, the mean and variance of the truncated normal are:

$$\begin{aligned} \boldsymbol{\mu}_{it,r} &= \mathbf{u}_i + W_{it,r} \theta + \Sigma'_{r,-r} \Sigma_{-r,-r}^{-1} (\mathbf{y}_{it,-r} - \mathbf{u}_i - W_{it,-r} \theta) \\ \Sigma_r &= \Sigma_{r,r} + \Sigma'_{r,-r} \Sigma_{-r,-r}^{-1} \Sigma_{r,-r} , \end{aligned} \quad (\text{A1.9})$$

where the dimension of  $\boldsymbol{\mu}_{it,r}$  is  $(r \times 1)$ ,  $\Sigma_r$  is dimension  $(r \times r)$ , and the indices  $r$  and  $-r$  refer to censored and positive elements, respectively (Huang, 2001). The fully augmented  $\mathbf{z}$  vector is subsequently used for drawing realizations of the parameters of interest from the conditional distributions for the model parameters.

The conditional posterior distributions are derived from specifications of prior distributions, which are specified using any previously known information about the parameters of interest. The prior distributions used in the random effects model for the parameters  $\theta$  and  $\Sigma$  are assumed independent and of the following form:

$$\pi(\theta) \sim N_K(\beta_0, B_0^{-1}) , \quad (\text{A1.10})$$

$$\pi(\Sigma) \sim IW_J(\rho_0, R_0) , \quad (\text{A1.11})$$

where  $\pi(\theta)$  is a  $K$ -dimension multivariate normal distribution with mean  $\beta_0$  and precision matrix  $B_0^{-1}$  and  $\pi(\Sigma)$  is a  $J$ -dimension inverse Wishart distribution with degrees of freedom  $\rho_0$  and scale  $R_0$ . The hyperparameters of the prior distributions  $(\beta_0, B_0^{-1}, \rho_0, R_0)$  are set to values that reflect very diffuse prior information. The values of  $\beta_0, B_0^{-1}, \rho_0,$  and  $R_0$  are set to 0,  $I_K, J,$  and  $I_J,$  respectively where  $I_K$  and  $I_J$  are  $K$ - and  $J$ -dimension identity matrices. With the values of the hyperparameters set, the conditional posterior densities of the model parameters are:

$$p(\theta|\mathbf{z}, \Sigma, \mathbf{u}, W) \sim N(\beta_1, B_1^{-1}), \quad (\text{A1.12})$$

$$p(\Sigma|\mathbf{z}, \theta, \mathbf{u}, W) \sim IW_J(\rho_1, R_1). \quad (\text{A1.13})$$

The posterior distribution of  $\theta$  is a  $K$ -dimension multivariate normal with mean

$$\beta_1 = \left( \sum_{i=1}^N \sum_{t=1}^{T_i} W_{it}' \Sigma^{-1} W_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^{T_i} W_{it}' \Sigma^{-1} \mathbf{d}_{it} \right), \text{ covariance matrix } B_1^{-1} = \left( \sum_{i=1}^N \sum_{t=1}^{T_i} W_{it}' \Sigma^{-1} W_{it} \right)^{-1}, \text{ and}$$

$\mathbf{d}_{it} = \mathbf{z}_{it} - \mathbf{u}_i$ . The posterior distribution of  $\Sigma$  is a  $J$ -dimension inverse Wishart with degrees of freedom  $\rho_1 = J + N * T$  and scale  $R_1 = (I_J J + \bar{S} N * T) / (J + N * T)$ , where

$$\bar{S} = \frac{1}{N * T} \sum_{i=1}^N \sum_{t=1}^{T_i} (\mathbf{d}_{it} - W_{it} \theta) (\mathbf{d}_{it} - W_{it} \theta)' \text{ and } T = \sum_{i=1}^N T_i.$$

In addition to these adjustments to the posterior distributions for  $\theta$  and  $\Sigma$ , the prior and posterior distributions of the random effects error components,  $\mathbf{u}_i$ , must be derived for the random effects model. The prior distributions for the error component,  $\mathbf{u}_i$ , and its variance,  $V$ , are assumed independent and of the following form:

$$\pi(\mathbf{u}) \sim N(\mu_0, M_0^{-1}), \quad (\text{A1.14})$$

$$\pi(V) \sim IW_J(\gamma_0, G_0), \quad (\text{A1.15})$$

where  $\pi(\mathbf{u})$  is a univariate normal distribution with mean  $\mu_0$  and precision matrix  $M_0^{-1}$  and  $\pi(V)$  is a  $J$ -dimension inverse Wishart distribution with degrees of freedom  $\gamma_0$  and scale  $G_0$ . As with the prior distributions of  $\theta$  and  $\Sigma$ , the hyperparameters are assumed to be known and are set to values that reflect very diffuse prior information. The values of  $\mu_0, M_0^{-1}, \gamma_0,$  and  $G_0$  are set to 0,  $V, J,$  and  $I_J,$  respectively where  $I_J$  is a  $J$ -dimension identity matrix. With these values of the hyperparameters, the posterior densities are derived as:

$$p(\mathbf{u}|\mathbf{z}, \theta, \Sigma, V, W) \sim N(\mu_1, M_1^2), \quad (\text{A1.16})$$

where  $\mu_1 = \left( \left( \sum_t \mathbf{z}_{it} - \sum_t W_{it} \theta \right) \Sigma^{-1} \right) M_1^2$  is the mean of the posterior and  $M_1^2 = (T I_J \Sigma^{-1} + V^{-1})^{-1}$  is

the variance. The posterior distribution of  $V$  is derived as follows:

$$p(V|\mathbf{z}, \theta, \Sigma, \mathbf{u}, W) \sim IW(\gamma_1, G_1), \quad (\text{A1.17})$$

where  $\gamma_1 = J + N$  are the degrees of freedom and  $G = (I_J J + \bar{S}N) / (J + N)$ , where  $\bar{S} = \frac{1}{N} \sum_{i=1}^N \mathbf{u}_i^2$  is the scale.

The following is an outline of the steps of the Gibbs sampler for estimation of the random effects model. The algorithm includes steps for sampling from the conditional distributions for both the household-specific error components and the variance of these errors. Iteration  $p$  of the Gibbs sampler algorithm is comprised of the following steps:

(1) Initialize the model unknowns with starting values,  $\theta^0, \Sigma^0, \mathbf{z}_{it}^0$ , where

$$z_{ijt}^0 = \begin{cases} y_{ijt} & \text{if } y_{ijt} > 0 \\ -1 & \text{if } y_{ijt} = 0 \end{cases} .$$

(2) At iteration  $p$ , complete the following:

- a. Draw realizations of  $\mathbf{y}_{it,r}^{*p} | \theta^{p-1}, \Sigma^{p-1}, \mathbf{u}_i, W_{it}, \mathbf{y}_{it,-r}$  for  $i=1, \dots, N$  from  $TN_{(-\infty, 0]}(\boldsymbol{\mu}_{it,r}^{p-1} + \Sigma_r^{p-1})$ , where  $\boldsymbol{\mu}_{it,r}$  and  $\Sigma_r$  are person specific as described above. Use the inversion method to draw from the truncated multivariate normal distribution given the most recent draws of the mean and variance of the distribution.
- b. Draw  $V^p | \mathbf{z}^p, \theta^{p-1}, \Sigma^{p-1}, \mathbf{u}^{p-1}, W$  from  $IW(\gamma_1, G_1)$ .
- c. Draw  $\mathbf{u}^p | \mathbf{z}^p, \theta^{p-1}, \Sigma^{p-1}, V^{p-1}, W$  from  $N(\boldsymbol{\mu}_1, M_1^2)$ .
- d. Draw  $\Sigma^p | \mathbf{z}^p, \theta^{p-1}, \mathbf{u}^p, W$  from  $IW_J(\rho_1, R_1)$ .
- e. Draw  $\theta^p | \mathbf{z}^p, \Sigma^p, \mathbf{u}^p, W$  from  $N_K(\boldsymbol{\beta}_1, B_1^{-1})$ .

(3) Repeat step (2) for  $p = 1, \dots, P$ , where  $P$  is large enough to obtain a sufficient number of posterior realizations.

## Appendix 2

### Bayesian Estimation: Convergence and Mixing

There are two primary concerns when implementing a Bayesian estimation methodology that uses a Markov Chain Monte Carlo (MCMC) algorithm: convergence and mixing. The MCMC algorithm must converge to the proper posterior density and should mix thoroughly across the support of that density (Lynch, pp.132-141, 2007). Trace plots of model parameters are useful for detecting convergence to the proper density. If the MCMC algorithm has not converged, trending will be seen in the trace plots. Trace plots for the commodity-specific media index and price variables of the random effects SUR tobit models are shown in figure A2.1. The trace plots for each parameter display a steady, stationary chain, indicating that convergence of the algorithm has been attained. The trace plots also appear to converge to the posterior density within about 20 iterations. Therefore, a burn in of 500 iterations is more than sufficient to make certain that posterior analysis is conducted using a converged model.

Histograms of the model parameters are also useful for diagnosing convergence and mixing. The histograms shown in figure A2.2 are the media index and price parameters of the random effects SUR tobit model. Recall that only every 36<sup>th</sup> posterior realization is kept in this model to decrease autocorrelation sufficiently. Therefore, the number of realizations that make up the histograms for the random effects model is 395. These histograms are approaching normal distributions, but are not sufficiently close to ensuring convergence and mixing. Therefore, a longer chain is needed to be confident in the results from the random effects model.

Another check of convergence for MCMC algorithm models is to begin the Gibbs sampler at different starting values. If the chains converge to the same posterior distribution, then the estimator is performing well. Figure A2.3 shows overlays of trace plots for the price parameters of the random effects models at different starting values. The starting values for Chain 1 (green) are the Ordinary Least Squares (OLS) estimates for the model. Chain 2 (blue) uses a starting value of 0.1 for each of the model parameters. The trace plots for each model indicate that convergence to the posterior distribution is robust to the selection of starting values.

Although MCMC algorithms such as the Gibbs sampler produce samples from a posterior distribution, these samples are not independent. The Markov property of the sampler uses the previous draw from the distribution as the basis for the next sample that is drawn. The samples are autocorrelated, which can cause the variance estimates to be incorrect. To account for autocorrelation between the samples in the chain, it is common to take every  $k^{\text{th}}$  draw for inference, where  $k$  is the lag beyond which autocorrelation no longer affects inference. The autocorrelation function (ACF) can be calculated to determine the appropriate number of sample to skip to have insignificant autocorrelation. The ACF for lag  $L$  is as follows:

$$\text{ACF}_L = \left( \frac{T}{T-L} \right) \frac{\sum_{t=1}^{T-L} (x_t - \bar{x})(x_{t+L} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2} , \quad (\text{A2.1})$$

where  $x_t$  is the sampled value of  $x$  for iteration  $t$ ,  $T$  is the total number of sampled values,  $\bar{x}$  is the mean of the sampled values, and  $L$  is the lag length (Lynch, pp.146-147, 2007).

The ACF was calculated for every parameter in the model. Figure A2.4 shows the ACF at different lag lengths for all the parameters in the random effects model. A lag length of 35 is required for all the parameters in this model to have an ACF of 0.25 or less. By omitting the first 500 iterations from the random effects model and keeping one in every 35 sampled values, 1,015 posterior realizations remain for inference.



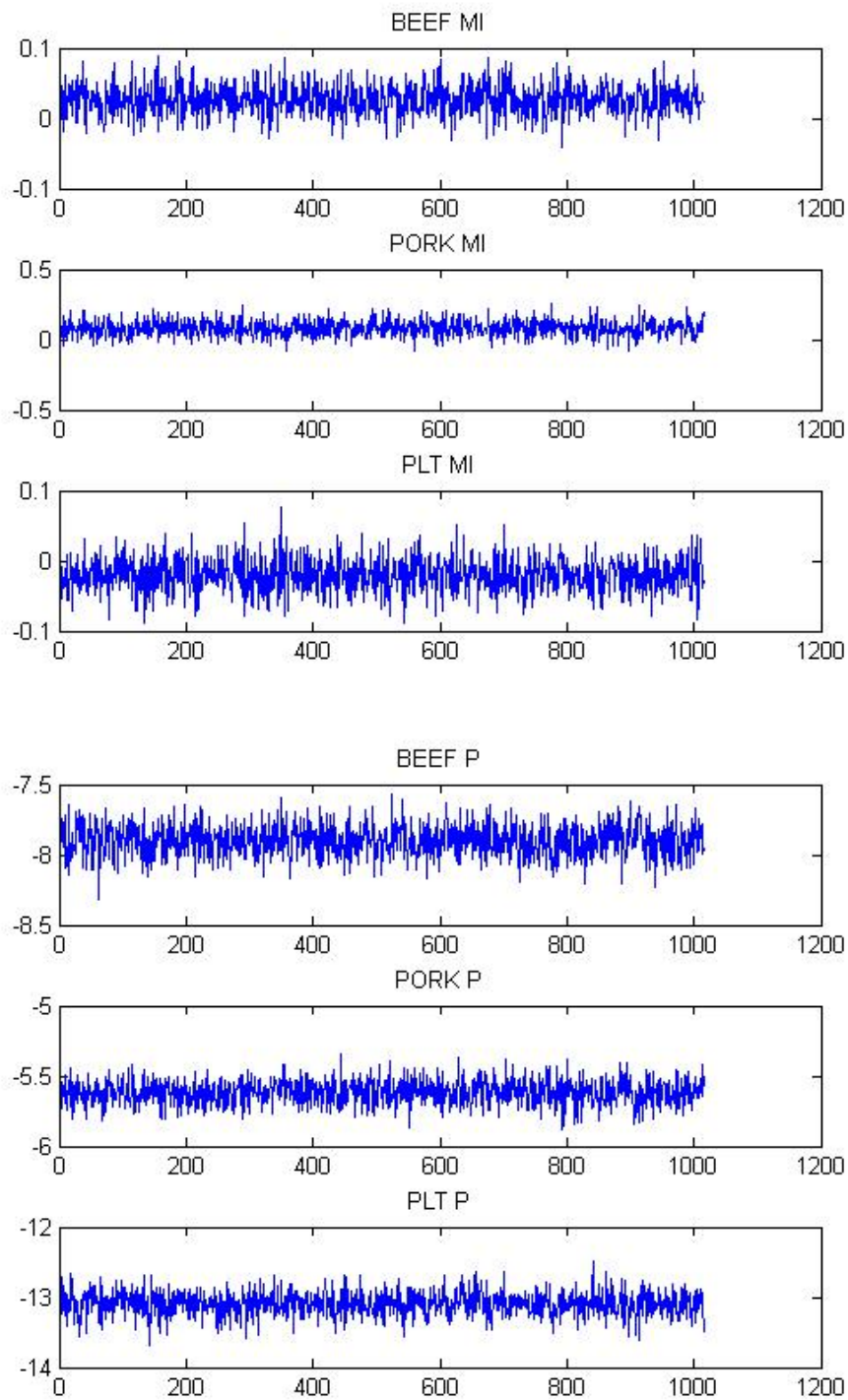


Figure A2.1 Trace Plots of Media Index and Price Parameters from Random Effects SUR Tobit Model

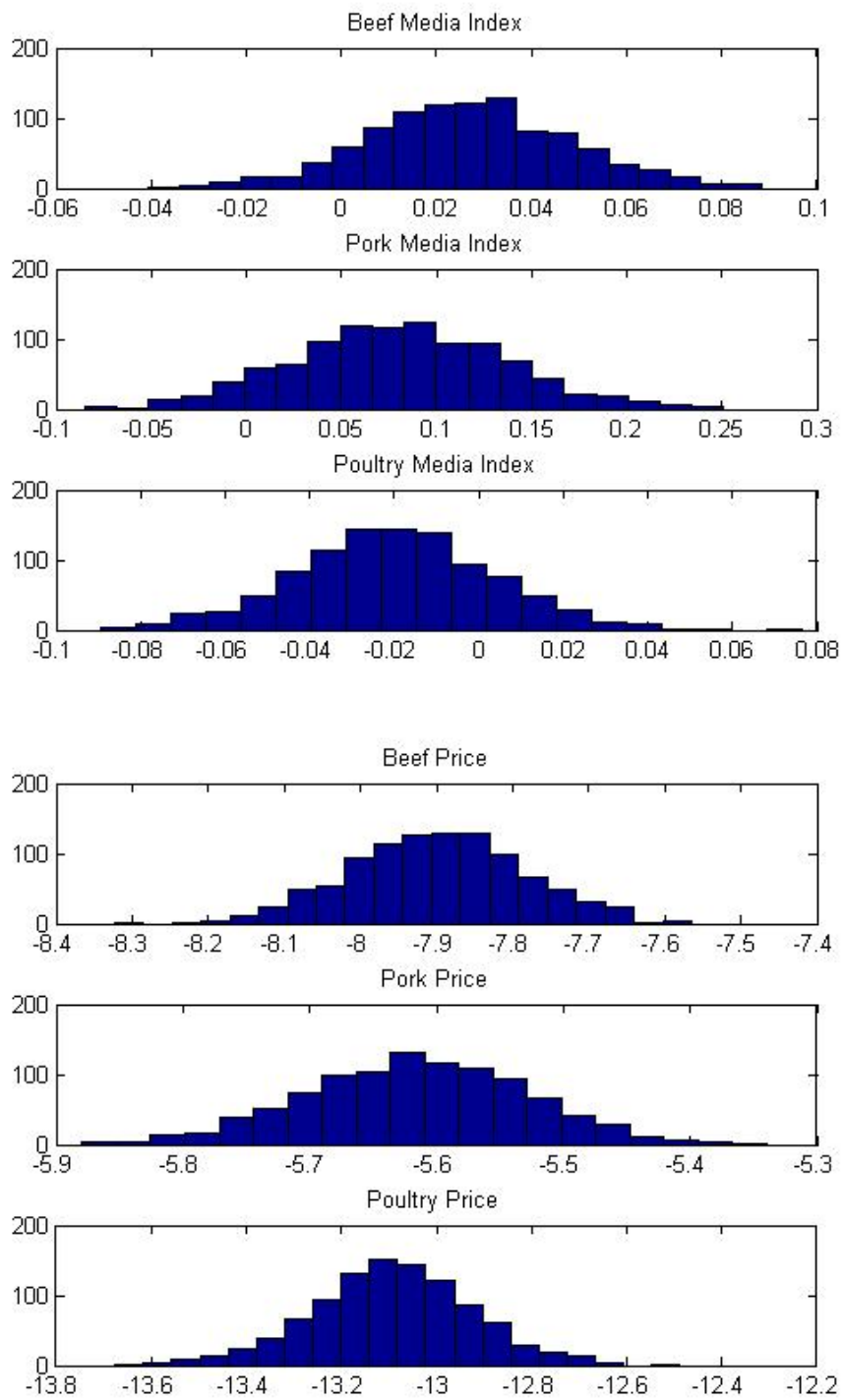


Figure A2.2 Histograms of Media Index and Price Parameters from Random Effects SUR Tobit Model

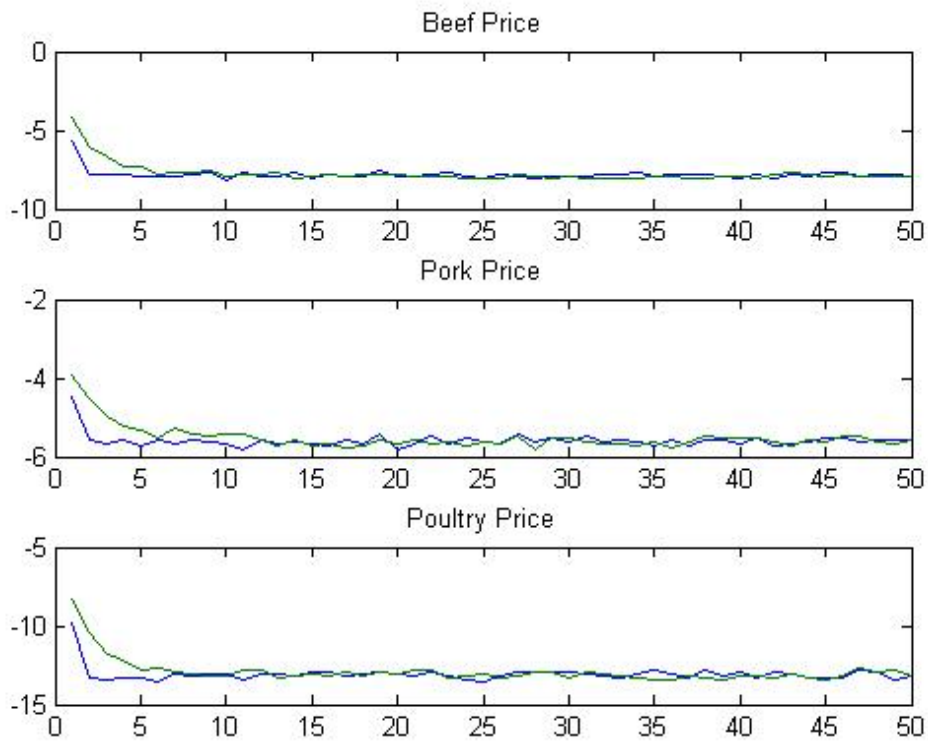


Figure A2.3 Trace Plots of Price Parameters from Random Effects SUR Tobit Model at Different Starting Values

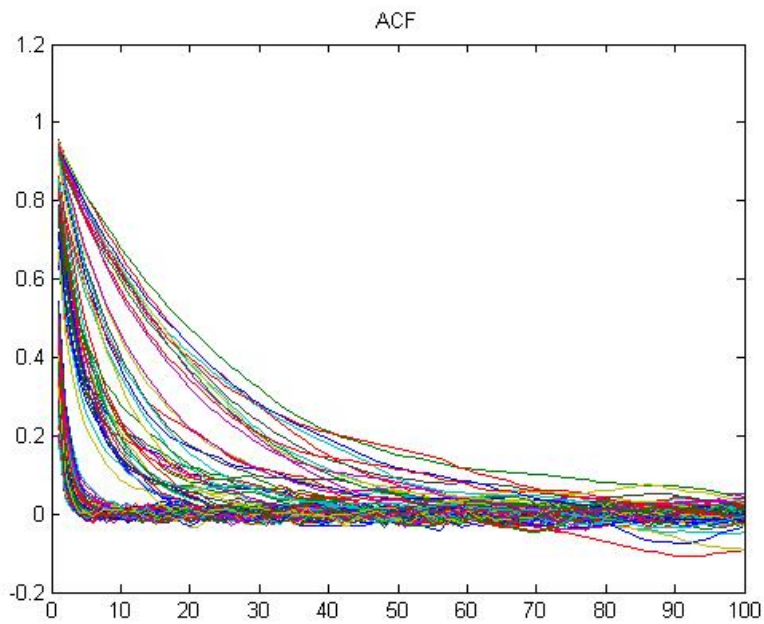


Figure A2.4 ACF at Different Lag Lengths of all the Parameters in the Random Effects SUR Tobit Model