School of Economic Sciences

Working Paper Series
WP 2010-17

Effects of Food Safety Information on Consumer Behavior: Evidence from U.S. Household-Level Data

By

Mykel R. Taylor

November 2010

WASHINGTON STATE UNIVERSITY
World Class. Face to Face.

Effects of Food Safety Information on Consumer Behavior: Evidence from U.S. Household-Level Data

by

Mykel R. Taylor[1]

School of Economic Sciences

Washington State University

---

[1] Assistant Professor: School of Economic Sciences, P.O. Box 646210, Washington State University, Pullman, WA 99164-6210. Phone: 509-335-8503. Fax: 509-335-1173. E-mail: m_taylor@wsu.edu

ABSTRACT

Many factors can influence consumer purchasing habits, including food safety information. While concerns about food safety are likely to be influenced by idiosyncratic experiences, general media information on the safety of meat and poultry may also affect purchase decisions. The reaction of consumers to changes in the amount of food safety information regarding fresh beef, pork, and poultry available in the media is the focus of this study. A discrete choice model is estimated to assess the probability that individual heterogeneous households will avoid making monthly meat and poultry purchases in response to changes in food safety information. Results suggest that some households do respond to changes in the level of publically available food safety information by choosing to avoid purchasing fresh meat or poultry.

JEL Classification: D12, Q18, D83

Keywords: food safety, discrete choice, information, consumer demand

For most consumers, the risks posed by foodborne pathogens amount to no more than a temporary case of 'food poisoning' with no lasting health effects. However, young children, the elderly, and people with compromised immune systems face an elevated risk of serious illness, or even death, if they come in contact with certain pathogens. As a result, every food recall notice posted by the Food Safety Inspection Service (FSIS) explicitly includes warnings targeted at these consumers. These warnings, along with recommendations from physicians and previous experience with foodborne illness, may affect the consumption decisions of households, regardless of their risk status.

While idiosyncratic experiences are difficult to measure, the amount of food safety information available to consumers in the press can be quantified. Food safety information can include product recalls by FSIS, educational information regarding good food safety practices at home, and opinion editorials regarding the overall safety of the U.S. food system, to name a few. The information varies in both the scope of food safety issues covered and whether or not the overall message is positive or negative, with regard to consumers risk exposure. This study uses a measure of media coverage of food safety information to mimic the variety of information sources and message content consumers are exposed to in their daily lives. The connection between food safety information and consumer purchase behavior is of interest to regulatory agencies, policy makers, and others because the use of recalls, and the subsequent media coverage of these recalls, is intended to change people's behavior. The efficacy of this type of information transfer can be measured, in part, by the extent to which consumers' modify their purchase decisions.

Previous research on consumer responses to food safety information has employed various measures of media coverage to infer its effect on food demand (e.g. Burton and Young, 1996; Piggott and Marsh, 2004). These studies have used aggregate data to jointly estimate meat and poultry demand equations that quantify the own- and cross-commodity effect of food safety information on marginal purchases. This approach has shown that media information matters at the aggregate level. A natural extension of this line of research is to use household-level meat and poultry purchase data to assess the likelihood that consumers will change their behavior in response to food safety information.

The objective of this article is to investigate if the quantity of food safety information that is publicly available impacts consumers' behavior with respect to the decision to purchase fresh meat and poultry. A media index is used as a proxy for food safety information available to consumers. The 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-specific, biweekly variables are constructed using the media index and a discrete choice model is estimated to measure the impact of food safety information on purchase behavior.

Results from this article suggest that peoples' response to food safety information does vary with certain household characteristics. Specifically, there is evidence that households with college educated or elderly heads, as well as households with children are more likely than other households to avoid purchasing meat and poultry when the level of food safety information in the media increases. These results suggest that public information regarding food safety influences not only targeted high-risk groups, but may also cause spill-over effects on college educated households that are not as likely to suffer a severe illness from foodborne pathogens.

Literature on Demand and Food Safety Information

The use of media indices to measure the impact of food safety information on demand has been employed in several aggregate-level demand studies. Smith, van Ravenswaay, and Thompson (1988) considered the effect of media publicity following a case of heptachlor contamination of fresh fluid milk in Hawaii on milk purchases. Significant negative effects on milk purchases were found from negative news coverage. However, positive news coverage did not appear to affect purchases, indicating that statements by the media assuring consumers of the safety of certain milk products were heavily discounted. Dahlgran and Fairchild (2002) studied the effect of adverse media coverage from salmonella contamination on the demand for chicken. Their model incorporated adverse media publicity from TV 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. Their results did indicate a negative demand response to adverse media, however, the effect died out in a matter of weeks. Unlike paid advertising, media coverage of food safety events can end abruptly as other news events take priority in programming. This lack of frequent message repetition was considered by the authors to be a possible reason for the absence of long-run alterations in demand.

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 model considered publicity on BSE to be a form of negative advertising and measured its effect using an index of media coverage. The index included both the number of articles per quarter and the cumulative number of articles to date for each quarter. BSE publicity was shown to have both significant short-run and long-run effects on consumer expenditures on beef and among the other meats with a decline in market share for beef of 4.5 percent by the end of 1993.

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. 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 beyond the contemporaneous quarter.

Other work focusing on food safety information and consumer demand has examined the impacts of specific events or information campaigns on consumer-level purchase behavior. Schlenker and Villas-Boas (2006) employed event study analysis to investigate the effects of media coverage of BSE on consumer and financial markets. They compared analysis results using three data sources: UPC-level scanner data, diary files from the U.S. Consumer Expenditure Survey (CES), and cattle futures prices. Statistically significant negative effects on purchases and cattle prices from media coverage of BSE were found using the UPC scanner and futures data. However, the CES did not reveal any statistically significant effect on consumer purchases or expenditures.

Using a reduced form analysis and household-level data from the U.S. Consumer Expenditure Survey, Shimshack, Ward, and Beatty (2007) examined responses to a U.S. national FDA advisory on exposure to methyl-mercury from store-bought fish. They employed both parametric and non-parametric methods to analyze changes in fish demand for households comprised of targeted consumers (i.e. households with young children, nursing mothers, and pregnant women) and non-targeted consumers. The analysis of subgroups of households in the sample revealed a short-run response to food safety information that was primarily determined by education level and newspaper readership. They also found spillover effects of decreased fish consumption among non-targeted households with high readership levels.

While previous studies have used both household-level data and media indices, none have used both types of data in the same empirical model. This approach was chosen to explicitly model household characteristics that could impact purchases, as well as to account for changes in consumer behavior over time, when repeatedly exposed to food safety events through media coverage. This article further expands the current literature by modeling consumer demand in a discrete choice framework. It is intuitively appealing to suggest that consumers who are concerned about or at high risk for serious illness from foodborne pathogens may make discrete changes in their purchase behavior under media coverage of a food safety event. These changes could manifest as substituting to other protein sources or abstaining completely for a period of time, which cannot be detected using a marginal demand model. Therefore, results reported in this article will expand our understanding of the impacts of food safety information on consumer choice in several ways.

Model of Meat and Poultry Purchases

Following notation found in Piggott and Marsh (2004), the household's utility function can be generally represented by $U(\mathbf{x}, \mathbf{r})$, where $\mathbf{x}$ is the quantity of fresh meat and poultry consumed and $\mathbf{r}$ is a vector of public information available to the consumer concerning the safety of meat and poultry. Rather than focusing on changes in the quantity of meat and poultry purchased, the consumer's decision to avoid meat and poultry entirely is modeled as a discrete decision to enter the market and make a purchase or remain out of the market for a given period of time.

The derivation of the model begins by specifying a random utility model where an individual, *n*, faces *J* alternatives. The utility a person gets from choosing one of the *J* alternatives is decomposed into an observed portion (i.e. known by the researcher), V_{nj}, and an unobserved portion, ε_{nj}, that is treated as random (Train, 2003). In this study, the observed components of the utility function include biweekly purchases of fresh meat and poultry, public information on food safety, seasonal fixed effects, and demographic characteristics specific to the individual household. Unobserved components of the utility function include, but are not limited to, previous experience with foodborne pathogens and personal health conditions that influence diet (e.g. high cholesterol, hypertension).

The utility of choosing a particular alternative is $U_{nj} = V_{nj} + \varepsilon_{nj}$, where ε_{nj} is distributed independently and identically as extreme value. Using Train's notation, the probability that individual *n* chooses alternative *j* is:

$$P_{nj} = \Pr\text{ob}\left(V_{nj} + \varepsilon_{nj} > V_{ni} + \varepsilon_{ni}, \forall j \neq i\right)$$
$$= \Pr\text{ob}\left(\varepsilon_{ni} < \varepsilon_{nj} + V_{nj} - V_{ni}, \forall j \neq i\right) . \quad (4)$$

The portion of utility that is observable, V_{nj}, is specified as a linear function of parameters as follows:

$$V_{nj} = \alpha_j + \boldsymbol{\beta}'_{nj}\mathbf{x}_{nj} \ , \quad (5)$$

where α_j is an alternative-specific constant term for alternative *j*, $\mathbf{x}_{nj}$ is a vector of containing both household- and alternative-varying characteristics, and the corresponding vector of estimated coefficients is $\boldsymbol{\beta}_{nj}$. If the utility of alternative *j* is greater than all other alternatives, then that will be the alternative that is chosen.

McFadden (1974) shows that if the error terms of the unobserved utility model are independent and identically distributed as Type I extreme value, then the probability of household *n* choosing any alternative *j* from *J* alternatives is:

$$P_{nj} = \frac{e^{V_{nj}}}{\sum_{j=1}^{J} e^{V_{nj}}} \ . \quad (6)$$

Estimation of this model requires that one of the *J* alternative-specific constants be normalized to zero. For the models described below, this is the 'no meat or poultry was purchased' option. Each of the alternative-specific constants are subsequently interpreted relative to this omitted option. The log likelihood function used in model estimation is as follows:

$$\ln_{\max \beta} L(\boldsymbol{\beta}) = \sum_{n=1}^{N} \sum_{j=1}^{J} \mathbf{d}_{nj} \ln P_{nj} \ , \tag{7}$$

where $\mathbf{d}_{nj}$ is an indicator vector with value equal to one if household *n* chose alternative *j* and zero otherwise.

The data used in this article allow for investigation of multiple product purchase patterns. For each two week period during the years 1998 to 2005, household purchases of beef, pork, and poultry are observed. Incorporating this information into a multinomial choice model allows for any interactions among the three commodities and reveals the probabilities of a household purchasing each of the goods as well as combinations of them. The eight purchase alternatives a household faces in a given two-week period are as follows: 1. beef; 2. pork; 3. poultry; 4. beef and pork; 5. beef and poultry; 6. pork and poultry; 7. beef, pork, and poultry; or 8. neither beef, pork, or poultry. Each household chooses one and only one of these alternatives.

The specification of the multinomial logit model follows the linear in parameters form shown in equation (5), which is comprised of parameters that vary across both alternatives and households. Using the media index as a proxy for food safety information, the model is specified as:

$$V_{njt} = \alpha_j + \sum_{k=1}^{3} \gamma_k Price_{nkt} I_k^j + \sum_{k=1}^{3} \sum_{l=1}^{3} \nu_{lk} Price_{nlt} (1 - I_k^j) + \sum_{k=1}^{3} \beta_k M_{nkt} I_k^j + \sum_{k=1}^{3} \sum_{l=1}^{3} \phi_{lk} M_{nlt} (1 - I_k^j) + \Gamma_{njt} (Z_n, M_{nkt}) + \Psi_{njt} + \Delta_{njt} \ , \tag{8}$$

where α_j is the j^{th} alternative specific constant, I_k^j is an indicator function that is equal to 1 if commodity $k \in$ the j^{th} alternative and equal to 0 otherwise, *k* and *l* each index the three commodities of interest, and *j* indexes the eight alternatives. The own-effect media index variable, M_{nkt}, is the interaction of the commodity- and region-specific media index variable for household *n* and the indicator function ($MI_{nkt} * I_k^j$). This variable is the value of the media index for commodity *k* if the indicator function equals 1 for commodity *l* and equal to 0 otherwise. The cross-effect media index parameter, M_{nlt}, is similarly defined. The matrix Γ_{njt} denotes interactions between household characteristics Z_n (age, education, children present, and urban location) and the own-effect media index M_{nkt}. The term Ψ_{njt} denotes other control variables such as time fixed effects and household characteristics (e.g. income, household size, race, geographic location). Finally, Δ_{njt} is a term that controls for household purchasing habits and inventory effects.

The variable *Price* used in the model is a 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*, uses a commodity- and region-specific media index that is based on the number of food safety articles appearing in U.S. regional newspapers in each two-week period. 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.

In addition to the household demographic and time fixed effects, variables measuring purchase decisions made in previous time periods are included in the model. These state-dependent variables capture both inventory and purchase habit effects. The variables are specified following Moeltner and Englin (2004) and consist of total numbers of purchases and total numbers of consecutive purchases. There are also corresponding totals for non-purchase and repeated non-purchase. It is expected that households that consistently purchase one of the commodities are likely to purchase again in the next period, whereas households that rarely purchase meat or poultry are unlikely to purchase in the next period. By explaining the variability due to state dependence, second-order effects from food safety information may be more accurately identified.

With the exception of the alternative-specific constants, the variables in this model are specified such that alternatives are 'bundled' into the commodities of beef, pork, and poultry. For example, rather than estimating a price coefficient for each of the eight alternatives, one price parameter is estimated for each of the three commodities. The estimated coefficient for the commodity-specific price coefficient, γ_k, is the effect of the price of commodity k on the probability of choosing an alternative that includes that commodity. The corresponding interpretation of the cross-price coefficient, ν_k, is the effect of the price of commodity l on the probability of choosing an alternative that includes commodity k. Similar interpretations are made for both the own-media index and the cross-media index variables.

To test the hypothesis that consumer response to food safety information may vary by household, interaction terms are included in the model for some of the household demographic characteristics. The interaction terms are specified between the food safety variable and the following four demographic variables: head of household with a college education or higher (*Ed*); head of household aged 55 or older (*Age*); location of the household in an urban area (*Urban*); and the presence of children in the household (*Child*). For example, the coefficient of the interaction term between the presence of children and the commodity-specific regional media index, ρ_k, would be interpreted as the effect of additional food safety articles pertaining to commodity *k* on the probability of purchasing commodity *k* for households with children present, relative to households without children. Interaction terms for the other demographic variables and the regional media index variable can be similarly interpreted.

Alternative-specific constants, α_j, are estimated for each alternative, except alternative 8 (no beef, pork, or poultry purchased) which is dropped from the model for estimation. These variables are not 'bundled' into commodity-specific coefficients, but rather are alternative-specific. The constants are interpreted as the average effect of non-included factors on the utility of an alternative relative to the omitted alternative of not purchasing beef, pork, or poultry. Variable definitions and summary statistics for the model variables are given in table 1.

Data

The multinomial logit demand model is specified using four different types of variables: household purchases of meat and poultry products, purchase prices, commodity-specific media indices, and household characteristics. The data needed to form these variables come from two sources. Data on household purchases and expenditures on meat and poultry products were obtained from the Nielsen Homescan panel. This panel covers households from all across the United States during the time period January 1998 to December 2005.[1] The Nielsen panel data also contain the information used to construct several demographic characteristics variables for 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 products of interest for this article 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. Each record 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.[2] A biweekly purchase periodicity was chosen for the empirical analysis to avoid excessive censoring rates, but still allow for short-run food safety effects. This frequency also reflects households' tendency to make meat and poultry purchases twice a month, which corresponds to the commonly used two-week pay period.[3]

Prices per unit of product were calculated by dividing total expenditure by total quantity for each individual meat or poultry purchase. 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 two-week period, 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.[4]

Imputation of the missing prices is based on the linear price model found in Cox and Wohlgenant (1986). The regression is specified using the average price of the good during a given time period from the consuming households in the panel. Household income is also 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 to capture non-linear effects of these variables. Other demographic variables were considered for the price equations, including region, race, and education, but are not used in the final specification of the price imputation model.

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

$$p_{itn} = \alpha \overline{p}_{it} + \boldsymbol{\gamma}_r \mathbf{r}_n + \delta u_n + \eta i_n + \kappa i_n^2 + \tau s_n + \rho s_n^2 + \varepsilon_{it} \ , \tag{1}$$

where p_{itn} is the observed price of good *i* in period *t* for consuming household *n*, $\overline{p}_{it}$ is the sample average biweekly price for good *i* in period *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, ε_{it} is an iid error term, and $\alpha, \gamma_r, \delta, \eta, \kappa, \tau,$ and ρ are the corresponding coefficients to be estimated.[5] 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 biweekly average prices and the geographic and demographic characteristics of the non-consuming households.

The grouping of purchases into various beef, pork, and poultry products (e.g. ground beef, roasts, bone-in and boneless pieces) having 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 biweekly 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, often unit prices are used to represent the average price. Unit prices, calculated by dividing total expenditures by total quantity purchased, can mask variation in product prices and quality. To avoid this problem, explicit consideration of variation within aggregate commodities is critical.

One way to account for the within-species price and quality variation that exists when purchases were aggregated is to use the group prices to create a price index. The Törnqvist (1936) price index used in this study is an expenditure share-weighted geometric price index.[6] It is a function of average prices and quantities of the beef, pork, and poultry groups, thereby controlling for individual product quality and price variation in the aggregation process. The commodity-specific price index is defined, for each household and time period, as follows:

$$p_{nt}^{B} = \prod_{i=1}^{G} p_{\text{int}}^{\;w_i} \;, \tag{2}$$

where p_{nt}^{B} is the index price of beef for household *n* in period *t*, p_{int} is the retail price of beef group *i* faced by the household *n* in period *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, which is based on average prices across all households and time periods, is calculated as follows:

$$w_i = \frac{\bar{p}_i \bar{x}_i}{\sum_{j=1}^{G} \bar{p}_j \bar{x}_j} \;, \tag{3}$$

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.[7] 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.

Following Piggott and Marsh (2004), food safety is measured using commodity-specific indices of newspaper articles. A commodity-specific index 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.[8] The article counts gathered from the regional newspaper search were linearly aggregated to create indices that are 15-day rolling averages of the number of newspapers articles published during the previous two weeks.[9] The intuition for this specification of the indices is that each day of the two week period is a potential purchase occasion and the available and relevant information for each purchase occasion may change as time passes. At the beginning of the two weeks, the articles most likely to impact household purchase decisions are the ones published in the latter half of the previous two weeks. Over the course of the two week period, however, the most relevant food safety information becomes articles published closer to the current two week period. The rolling average specification captures this change in available information over the two week period. Figures 1-3 display the regional media indices for each of the three commodity groups.

Empirical Model

The model specified in equation (8) includes interaction terms between the own-effect media index variable and select demographic variables to test if food safety information impacts on consumer behavior differ across households. The education variable 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.[10] 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 being present in the household 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 interacted with food safety information is a binary variable indicating if the household is located in an urban area. It equals one if the household resides in an urban area and equals zero otherwise.

The intuition for including food safety interaction terms with the demographic variables for children and head of household aged 55 and older in the model is that 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 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 also enter the model separately to account for the average effects of these characteristics.

Other variables included in the model are household specific and intended to account for variability in purchase behavior that does not stem from food safety information. They include variables for household income and a quadratic household income term. 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.[11] The size of the household is also included in the regression to account for possible differences in purchase patterns for large versus small families. Seasonal effects in the purchase patterns of households are accounted for using monthly dummy variables with the parameter for December omitted from the regression. Annual effects in demand are also considered using year dummy variables with the variable for 2003 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 with the variable for the southern region dropped from the regression. The race of the head of household is categorized by Caucasian, Hispanic, black, Asian, and other. The race variable Caucasian is omitted from the regression. The expected signs of the geographic location and race variables are not known a priori.

The size of the full Nielsen dataset is 1,604,746 biweekly observations. A sample of this size presents a number of challenges to estimation. Therefore, is was determined that a bootstrap estimation method, using sub-samples from the original dataset would be appropriate. The model is estimated by drawing without replacement 500 sub-samples of 1,000 unique households, from the full data sample. This procedure prevents the model results from being influenced by any one sample from the panel dataset.

The estimated coefficients are calculated from the bootstrap dataset with B rows corresponding to the number of sample replications (500 in this study) and K columns corresponding to the number of variables in the model. Parameter estimates are calculated as follows:

$$\bar{\beta} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b \quad , \tag{9}$$

where $\hat{\beta}_b$ is the estimated parameter from the *b*th replication. To compute the variance of $\bar{\beta}$, note the following:

$$Var\left(\bar{\beta}\right) = \frac{1}{B^2}\operatorname{var}\left[\sum_{b=1}^{B}\hat{\beta}_b\right] \tag{10}$$

$$= \frac{1}{B^2}\sum_{b=1}^{B}\operatorname{var}\left(\hat{\beta}_b\right)$$

$$= \frac{1}{B}\operatorname{var}\left(\hat{\beta}_b\right) \ .$$

Equation (10) is a result of bootstrap replications being independent, by construction, and the distribution of $\hat{\beta}_b$ being identical for all *b*.

Results

The parameter estimates for the variables measuring price, food safety, and the interaction terms for select households, are given in table 2 (estimates of the full model are given in table A1 of the appendix). The coefficients are interpreted as the effect of a given factor on the probability a household will make a purchase of either beef, pork, or poultry. A positive sign indicates an increase in the probability, while a negative sign indicates a decrease in the probability of purchase. Coefficients are statistically significantly different from zero using a 95% confidence interval. The t-statistic for this confidence interval is approximated using the standard error of the distribution of the 500 bootstrapped parameter estimates.

The results indicate the price coefficients for beef, pork, and poultry all have the expected negative sign and are statistically significantly different from zero. The negative signs of all the price coefficients indicate that an increase in the price of any of the three meat commodities will decrease the likelihood of purchase, relative to purchasing no meat or poultry at all. Most of the cross-price coefficients are not statistically different from zero at the 5 percent level. The two cross-price coefficients that are statistically significant are the effects of beef and poultry price on the probability of purchasing pork. Both of these coefficients have a positive sign, indicating that an increase in the price of beef or poultry will increase the probability of making a purchase that includes pork. The signs of the cross-price coefficients suggest that consumers consider beef, pork, and poultry to be price substitutes with regard to the decision to make a purchase.

The multinomial logit model results indicate that food safety information can have a statistically significant impact on the probability of purchasing fresh meat and poultry for certain households. The total marginal effect of food safety information for each type of household considered is the sum of the own-effect and the interaction effect.[12] These total effects are listed in table 3 for the sum of the average media index parameters and each of the demographic interaction parameters.

The own-effects of food safety information are positive and statistically significant for both beef and pork. For some households, the interaction coefficients are larger in magnitude and have an offsetting negative sign, as compared to the own-effect coefficients. A total negative effect for high education households suggests that these households are less likely to buy beef and poultry when the amount of food safety information increases. This negative effect also occurs for purchases of beef and poultry by elderly households, pork and poultry purchases for households with children, and poultry purchases by urban households. A negative total marginal effect for a given commodity indicates that certain households are more likely to abstain from making a purchase within this fresh meat or poultry category. This is a striking result given that most recalls are associated with a single product within the fresh meat or poultry category (e.g. ground beef). The negative effect implies that there is a higher probability some consumers will avoid purchases of the entire category and not choose to substitute for other products within the category, such as steaks or roasts.

The total marginal effect on beef and pork is positive for some households. College educated households, elderly households, and urban households are more likely to purchase pork when the food safety media index increases. Households with children and those located in urban areas are also more likely to make a beef purchase when the media index increases. While the signs of each total marginal effect are either negative or positive, the magnitude is quite small for some of the parameters and, therefore, not likely to be different from zero.

A zero or positive total marginal effect for a given commodity and type of household may be interpreted in more than one way. First, if the effect is not different from zero, this could imply that households are still choosing to buy products within a given fresh meat or poultry category. They may simply be avoiding the recalled product.

For example, if a recall is issued for ground beef, households may choose to buy roasts or steaks and their probability of making a purchase within the category is unaffected by changes in the media index. Another possibility is that households may infer that the recall for ground beef justifies throwing out their current inventory and going to the store to replenish their stock. This behavior could explain a total positive effect on the probability of purchase when the level of the food safety media index increases.

While the interpretation of the negative total marginal effects appears relatively straightforward, the zero and positive signs on some parameters suggest that the data used to estimate this model may not be sufficiently disaggregated. A model using product-level, rather than commodity-level, groups could provide a clearer explanation for the zero and positive signs and subsequent inferences about consumer behavior. For example, the subgroups used in this analysis include both ground beef and boneless poultry. From a consumer standpoint, it seems likely that these two products have similar uses in at-home cooking and would be closer substitutes than some of the other subgroups. This type of cross-commodity substitution is not directly investigated in the current analysis and remains an important issue for future research.

It is worth noting that the state dependent variables included in the model were statistically significant. The variables measuring total number of past purchases and total number of consecutive past purchases indicate a positive effect on the probability of purchase in the current period. Correspondingly, the variables measuring non-purchase indicate that as the total number of periods (or total number of consecutive periods) where no meat or poultry was purchased increase, the probability of purchase declines.

The statistical significance of the state dependent variables suggests that discrete choice model specifications that do not account for habit or inventory effects may be incorrectly attributing consumer behavior to other factors. Further investigation of the state dependence variables and their interaction with other factors is left for future research.

Conclusion

The objective of this study was to investigate if the quantity of food safety information available to consumers impacts their behavior through purchase decisions for fresh meat and poultry in a discrete choice framework. Specifically, the demand model was designed to determine if food safety information effects vary across heterogeneous households. The model was estimated using a discrete choice framework, which allows for households that abstain from purchasing meat and poultry when the amount of food safety information available to the public increases.

The parameter estimates of the multinomial logit model suggest that there are differences in household-level responses to food safety information and that avoidance behavior is detectable for certain households. Each of the household types used in the model are likely to stop purchasing poultry in a given two-week period, when the amount of food safety information increases. This effect can also be found for beef with college educated and elderly households, as well as pork purchases for households with children. The probability of purchase for other product-household combinations considered in the model is either zero or positive.

A negative effect on the probability of purchase by households with elderly heads and children in response to more food safety information is not surprising, given the amount of targeted information that is disseminated from government agencies to these high risk groups. The response by college educated households is interesting, given that they are not necessarily at risk for serious illness from foodborne pathogens. This finding is similar to the results reported by Shimshack, Ward, and Beatty's (2007) study on FDA advisories of mercury contamination in fish. The results indicates that food safety information targeted toward some consumers can 'spillover' and become a factor in purchase decisions for both high and low risk consumers.

The results of this demand analysis also suggest that the aggregate demand models used in previous studies are not well suited to describing the full distribution of consumer response to food safety information. The results of the Piggott and Marsh (2004) study suggest that consumer reaction to a major food safety event tends to be relatively small in magnitude and short-lived. However, by analyzing the impacts on individual households, it was possible to find evidence that the reaction of certain households may be quite different as compared to the mean response. These results provide reasonable assurance that food safety warnings and information provided by FSIS, the media, or others can contribute to a reduction in exposure of high-risk groups to serious illness or death from foodborne pathogens.

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.

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.

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.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, ed. P. Zarembka, 105-142. New York: Academic Press, 1974.

Moeltner, K. and J. Englin. "Choice Behavior Under Time-Variant Quality: State Dependence Versus "Play-It-By-Ear" in Selecting Ski Resorts." *Journal of Business and Economic Statistics.* 22(2004): 214-224.

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/

Shimshack, J.P., M.B. Ward, and T.K.M. Beatty. "Mercury Advisories: Information, Education, and Fish Consumption." *Journal of Environmental Economics and Management*. 53(2007): 158-179.

Smith, M., E. van Ravenswaay, and S. Thompson. "Sales Loss Determination in Food Contamination Incidents: An Application to Milk Bans in Hawaii." *American Journal of Agricultural Economics*. 70(1988): 513-520.

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.

White, H. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." *Econometrica*, 48(1980):817-838.

Table 1 Summary Statistics of Demand Model Variables

	Description	Average	Minimum	Maximum	Std. Dev.
Beef Price	Quality-adjusted, beef price ($/lb)	3.209	0.577	12.638	0.562
Pork Price	Quality-adjusted, pork price ($/lb)	2.534	0.627	12.219	0.509
Poultry Price	Quality-adjusted, poultry price ($/lb)	1.924	0.700	8.195	0.248
Beef MI	Beef regional media index	7.633	0.786	77.645	6.428
Pork MI	Pork regional media index	2.547	0	16.567	1.988
Poultry MI	Poultry regional media index	11.378	2.000	38.310	6.054
Ed	= 1 if head of household is college educated	0.393	0	1	0.488
Age	= 1 if head of household is aged 55 or older	0.372	0	1	0.483
Urban	= 1 if household is in urban location	0.875	0	1	0.330
Child	= 1 if household has children present	0.296	0	1	0.456
Income	Total household income ($10,000)	5.383	0.250	12.500	3.151
Income^2	Total household income ($10,000), squared	38.910	0.062	156.250	43.477
Hsize	Total number of persons in household	2.532	1	9	1.379
Y1	= 1 if year is 1998	0.120	0	1	0.325
Y2	= 1 if year is 1999	0.112	0	1	0.316
Y3	= 1 if year is 2000	0.118	0	1	0.322
Y4	= 1 if year is 2001	0.127	0	1	0.333
Y5	= 1 if year is 2002	0.133	0	1	0.340
Y6	= 1 if year is 2003	0.136	0	1	0.342
Y7	= 1 if year is 2004	0.129	0	1	0.336
Y8	= 1 if year is 2005	0.125	0	1	0.330
M1 - M12	Monthly binary variables for Jan-Dec	0.083	0	1	0.276
South	= 1 if household located in southern region	0.366	0	1	0.482
Central	= 1 if household located in central region	0.204	0	1	0.403
West	= 1 if household located in western region	0.217	0	1	0.412
Northeast	= 1 if household located in northeastern region	0.213	0	1	0.410
Caucasian	= 1 if race of head of household is caucasian	0.766	0	1	0.423
Hispanic	= 1 if race of head of household is Hispanic	0.076	0	1	0.264
Black	= 1 if race of head of household is black	0.121	0	1	0.326
Asian	= 1 if race of head of household is Asian	0.022	0	1	0.146
Other	= 1 if race of head of household is other	0.016	0	1	0.126

Note: The number of observations is 745,632.

Table 2 Estimated Coefficients of Select Demand Model Parameters

		Bootstrap Sample		
	Alternative	**Coefficent**	**Standard Error**	**t statistic**
Price - Own	Beef	-1.143	0.007	-169.356
	Pork	-1.288	0.008	-159.253
	Poultry	-3.025	0.014	-216.064
Price - Beef	Pork	0.194	0.002	92.771
	Poultry	0.321	0.004	83.698
Price - Pork	Beef	0.086	0.002	35.194
	Poultry	0.038	0.004	9.019
Price - Poultry	Beef	0.058	0.002	27.538
	Pork	0.106	0.002	51.454
MI - Own	Beef	0.001	1.9E-04	2.918
	Pork	0.003	0.001	4.533
	Poultry	-7.8E-05	2.9E-04	-0.269
MI - Beef	Pork	0.003	2.0E-04	13.089
	Poultry	-0.003	7.7E-05	-41.620
MI - Pork	Beef	0.001	8.3E-05	10.401
	Poultry	-0.003	9.4E-05	-31.629
MI - Poultry	Beef	-0.001	6.9E-05	-14.316
	Pork	0.007	2.1E-04	32.451
Ed*MI	Beef	-0.002	1.2E-04	-15.892
	Pork	-0.003	4.4E-04	-6.269
	Poultry	-4.1E-04	1.6E-04	-2.636
Age*MI	Beef	-0.001	1.3E-04	-7.444
	Pork	0.001	4.9E-04	1.988
	Poultry	-0.001	1.8E-04	-4.923
Child*MI	Beef	0.000	1.4E-04	0.037
	Pork	-0.005	4.9E-04	-9.172
	Poultry	-0.002	1.9E-04	-9.052
Urban*MI	Beef	0.001	1.6E-04	5.196
	Pork	0.001	0.001	1.753
	Poultry	-3.9E-04	0.000	-1.489

Table 3 Estimated Total Marginal Effect of Select Demand Model Parameters

	Alternative	Commodity Average Effect	Demographic Interaction Effect	Total Effect
Ed*MI	Beef	0.001*	-0.002*	-0.001
	Pork	0.003*	-0.003*	2.7E-04
	Poultry	-7.8E-05	-4.1E-04*	-4.9E-04
Age*MI	Beef	0.001*	-0.001*	-4.4E-04
	Pork	0.003*	0.001*	0.004
	Poultry	-7.8E-05	-0.001*	-0.001
Child*MI	Beef	0.001*	5.1E-06	0.001
	Pork	0.003*	-0.005*	-0.002
	Poultry	-7.8E-05	-0.002*	-0.002
Urban*MI	Beef	0.001*	0.001*	0.001
	Pork	0.003*	0.001	0.004
	Poultry	-7.8E-05	-3.9E-04	-4.7E-04

Note: Coefficient estimates denoted with a * are statistically significant using a 95% confidence interval. Statistical significance of the individual parameter estimates does not imply statistical significance of the total marginal effect.

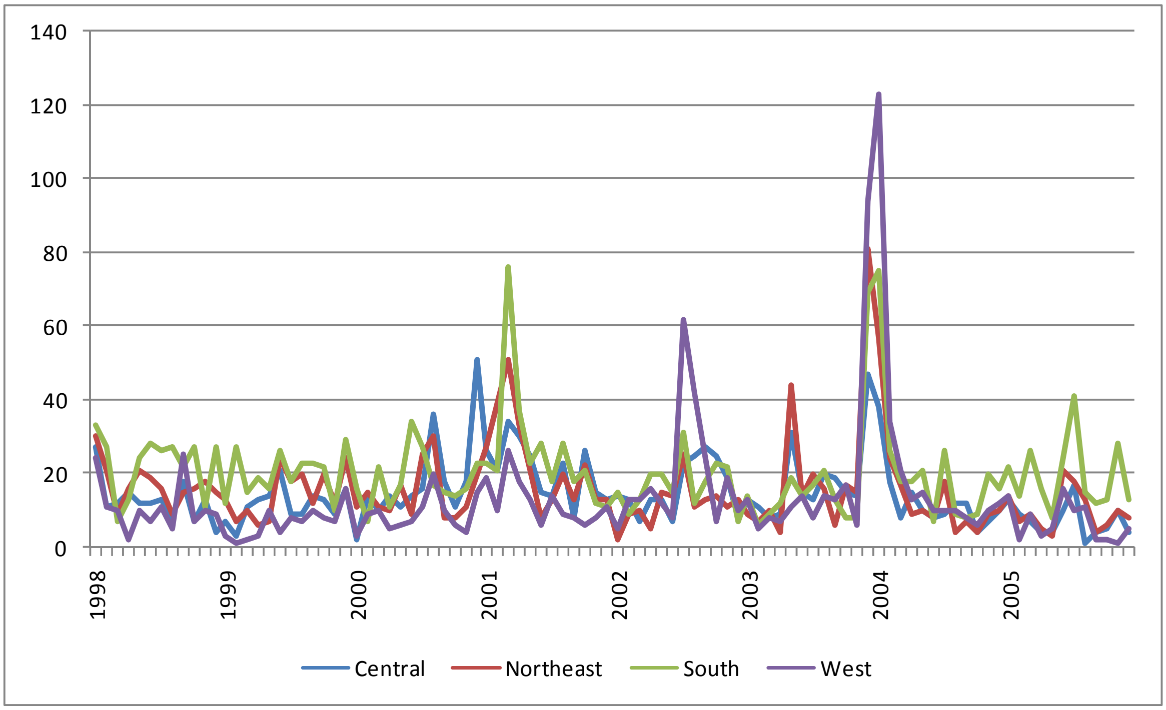

Figure 1 Beef Media Index by Region, 1998 to 2005

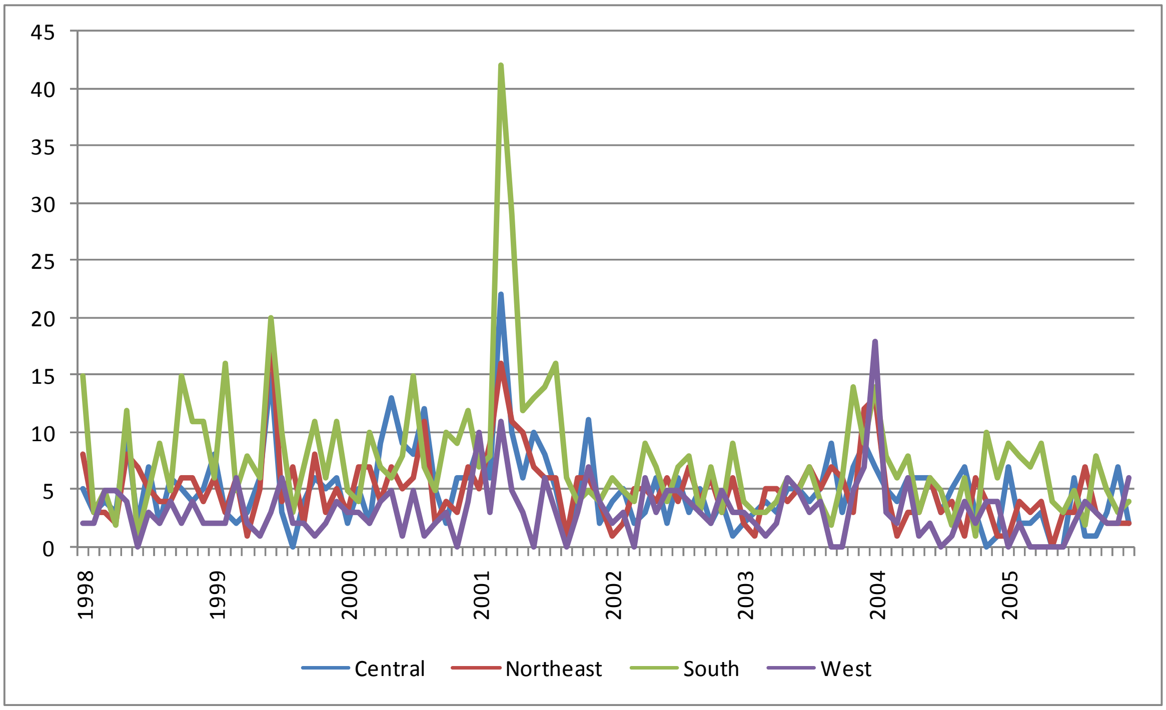

Figure 2 Pork Media Index by Region, 1998 to 2005

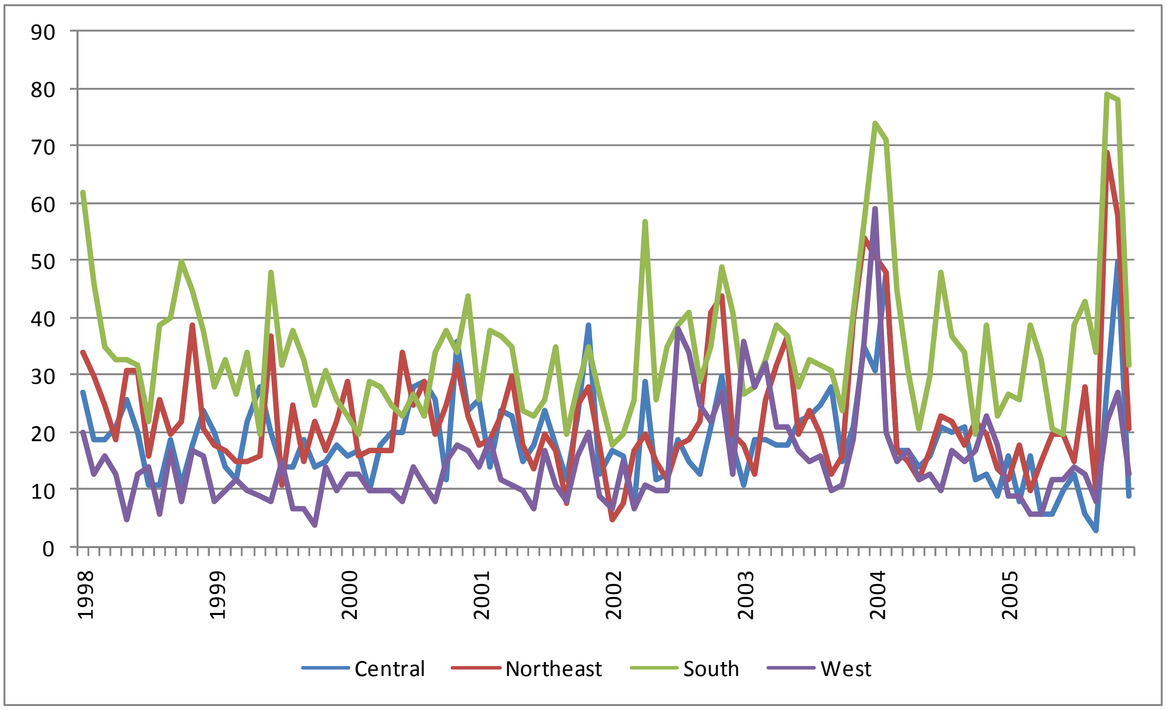

Figure 3 Poultry Media Index by Region, 1998 to 2005

[1] 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. Data include detailed product information, date of purchase, total quantity, total expenditure, and the value of any coupons used for every item purchased. The household sample is selected to correspond with the U.S. Census demographic distribution.

[2] If multiple purchases were made on a given day, each purchase is recorded as a separate observation in the raw dataset.

[3] Earlier models estimated using a monthly time period did not reveal any change in behavior with regard to probability of purchase. Given the lack of statistically significant parameter estimates in a monthly model and a biweekly average number of shopping trips recorded in this panel for fresh meat and poultry purchases, it is important to consider time periods in the data aggregation that correspond to observed behavior.

[4] The proportion of observed versus predicted prices varied substantially across the subcategories of products that comprised the beef, pork, and poultry groups. On average the proportion of observed transaction prices to the total number of transaction prices was 19.7 percent, 11.1 percent, and 11.0 percent for beef, pork, and poultry, respectively.

[5] 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.

[6] Motivation for using the Törnqvist price index stems from Diewert's (1976) paper on index number theory.

[7] The biweekly retail price of each group is the observed group price if the household bought that group in period t. If the household did not purchase that group, then the predicted group price is used.

[8] The article queries were constructed using the keywords *food safety* or *contamination* or *product recall* or *outbreak* or *salmonella* or *listeria* or *E. coli* or *trichinae* or *staphylococcus* or *foodborne*. From these search results, the articles were further queried for commodity-specific information using the search terms *beef* or *hamburger*; *pork* or *ham*; and *chicken*, *turkey*, or *poultry*.

[9] 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.

[10] 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.

[11] 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.

[12] Statistical significance of the individual coefficients for the own- and interaction effects does not necessarily imply statistical significance of the total marginal effect.