ABSTRACT

TAYLOR, MYKEL R. The Impact of Food Safety Information on Consumer Demand for Meat and Poultry: Evidence of Heterogeneous Household Effects. (Under the direction of Professor N.E. Piggott.)

Many factors can influence consumer purchasing habits, including food safety information. 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. However, 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. The reaction of consumers to changes in the amount of food safety information on beef, pork, and poultry available in the media is the focus of this study. Specifically, any differences in consumer reactions due to heterogeneous household characteristics are investigated.

Consumer reactions are modeled using both discrete and continuous choice models. Discrete choice models are 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 of a multinomial logit model suggest that some households do respond to changes in the level of food safety information available by choosing to avoid purchasing meat or poultry. Purchase avoidance behavior is also analyzed with a discrete-continuous model that employs monthly household-level panel data. A seemingly unrelated regression (SUR) tobit model is estimated using a Gibbs sampler with data augmentation. A component error structure is used to model unobserved heterogeneity of households making

repeated purchases over time. Food safety elasticities calculated from the random effects SUR tobit model do not provide much evidence that food safety information has an economically significant effect on household purchases of meat and poultry.

The Impact of Food Safety Information on Consumer Demand for Meat and Poultry:
Evidence of Heterogeneous Household Effects

by
Mykel Rae Taylor

A dissertation submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina

July 3, 2008

APPROVED BY:

Dr. Nicholas E. Piggott Committee Chair	Dr. Daniel Phaneuf
Dr. Michael Wohlgenant	Dr. Roger von Haefen

DEDICATION

This dissertation and the past four years of work are dedicated to Justin and Camille.

BIOGRAPHY

Mykel R. Taylor was born and raised on a cattle ranch in Montana. After graduation from Roberts High School, she traveled in Up With People across the United States, Canada, and northern Europe. She graduated with a B.S. in Agribusiness in 2000 and a M.S. in Applied Economics in 2001 from Montana State University. Mykel spent the next two years working at Kansas State University as an Extension Assistant in the Department of Agricultural Economics. She began her education at North Carolina State University in the fall of 2004.

ACKNOWLEDGEMENTS

I would like to acknowledge my committee for their assistance in the research and writing process. Their insight and guidance has given me the confidence to proceed in the field of economics as a researcher and educator. I would also like to thank my family and friends for their support throughout my graduate career. They are the reason I was able to come to work every day.

TABLE OF CONTENTS

LIST OF TABLES viii

LIST OF FIGURES x

1 Overview and Objectives of Research 1
 1.1 Overview 1
 1.2 Objectives and Contributions 2
 1.3 Outline of Dissertation 3

2 Literature Review 5
 2.1 Demand and Food Safety Information 5
 2.2 Demand Studies using Microeconomic Data 9
  2.2.1 Censored Observations and Demand System Estimation 10
  2.2.2 Price-Quality Variation 12
  2.2.3 Quantity Aggregation 14
 2.3 Conclusion 15

3 Meat and Poultry Consumption and Food Safety Data 19
 3.1 Introduction 19
 3.2 U.S. Household Consumption Data 19
  3.2.1 Meat and Poultry Purchases 20
  3.2.2 Household Demographics 24
  3.2.3 Demographics and Purchases of Select Groups 25
 3.3 Missing Prices 29
 3.4 Price Indices 32
 3.5 Quality Adjusted Quantities 33
 3.6 Food Safety Media Indices 34
 3.7 Conclusion 36

4 Discrete Choice Models of Meat and Poultry Purchases 58
 4.1 Introduction 58
 4.2 Binary Choice Models 59

4.2.1 Logit Model Derivation 60
4.2.2 Model Specification 62
4.2.3 Estimation and Results 65
4.2.3.1 Beef Purchase Models 66
4.2.3.2 Pork Purchase Models 68
4.2.3.3 Poultry Purchase Models 70
4.2.4 Binary Choice Model Elasticities 72
4.2.5 Binary Choice Model Summary 75
4.3 Multiple Choice Models of Meat and Poultry Purchases 75
4.3.1 Alternative-Specific Logit Model Derivation 76
4.3.2 Model Specification 78
4.3.3 Estimation and Results 82
4.3.4 Multinomial Logit Elasticities 86
4.3.5 8-Choice Model Summary 88
4.4 Conclusion 89
5 Demand Models of Meat and Poultry Consumption 102
5.1 Introduction 102
5.2 Demand System Model 103
5.2.1 Pooled SUR Tobit 104
5.2.2 Random Effects SUR Tobit 106
5.3 Estimation Methodology 109
5.3.1 Bayesian Estimation of the Pooled SUR Tobit Model 112
5.3.2 Bayesian Estimation of the Pooled SUR Tobit Model 117
5.3.3 Demand Model Specification 123
5.4 Results 125
5.4.1 Posterior Samples and Convergence 126
5.4.2 Pooled SUR Tobit Model Parameter Estimates 128
5.4.3 Random Effects SUR Tobit Model Parameter Estimates 131
5.4.4 Elasticities 134
5.5 Conclusion 141
6 Conclusions and Future Research 152

6.1 Conclusion 152
6.1.1 Discrete Choice Models of Meat and Poultry Purchases 153
6.1.2 Demand Models of Meat and Poultry Consumption 155
6.2 Future Research 156
6.2.1 Model and Variable Specification 157
6.2.2 Alternative Estimation Techniques 158
6.2.3 Structural Demand Models 159
Bibliography 162
Appendices 167
A Bayesian Estimation Issues 168
A.1 Convergence and Mixing 168

LIST OF TABLES

Table 2.1 Summary of Demand and Food Safety Literature 16

Table 2.2 Summary of Demand Research using Microeconomic Data 18

Table 3.1 Summary Statistics of Household Meat and Poultry Purchases 39

Table 3.2 Censored Monthly Observations by Year 41

Table 3.3 Censored Monthly Observations by Commodity 41

Table 3.4 Household Participation Rates 41

Table 3.5 Household Panel Demographic Variables 42

Table 3.6 Household Participation Rates by Demographic Group 44

Table 3.7 Household Panel Demographic Variables by Demographic Group 45

Table 3.8 Monthly Meat and Poultry Purchases by Demographic Group 47

Table 3.9 Estimated Coefficients of Censored Product Price Models 48

Table 3.10 Summary Statistics of Observed and Predicted Monthly Prices 50

Table 3.11 Summary Statistics of Quality Adjusted Monthly Purchases and Price Indices .....

.... 51

Table 3.12 Names and Locations of Newspapers used in Media Index 51

Table 3.13 Summary Statistics of Monthly Household Purchase Occasions 51

Table 3.14 Summary Statistics of Monthly Food Safety Information, 1998 to 2005 52

Table 4.1 Summary Statistics of Binary Choice Model Variables 92

Table 4.2 Estimated Coefficients of Binary Choice Model Variables 93

Table 4.3 Binary Choice Model Elasticities 95

Table 4.4 Summary Statistics of 8-Choice Model Variables 96

Table 4.5 Estimated Coefficients of 8-Choice Model 98

Table 4.6 8-Choice Model Elasticities 101

Table 5.1 Summary Statistics of Demand Model Variables 144

Table 5.2 Bayesian Estimated Coefficients of the Pooled SUR Tobit Model 145

Table 5.3 Bayesian Estimated Coefficients of the Random Effects SUR Tobit Model 147

Table 5.4 Price and Food Safety Elasticities of Pooled SUR Tobit Model 149

Table 5.5 Price and Food Safety Elasticities of Random Effects SUR Tobit Model 150

LIST OF FIGURES

Figure 3.1 Nielsen Panel Household Characteristics 53

Figure 3.2 Nielsen Panel Household Characteristics by Demographic Group 54

Figure 3.3 Frequencies of Monthly Household Purchase Occasions for Beef, Pork, and Poultry Products 55

Figure 3.4 Beef Media Index by Region, 1998 to 2005 55

Figure 3.5 Pork Media Index by Region, 1998 to 2005 56

Figure 3.6 Poultry Media Index by Region, 1998 to 2005 56

Figure 3.7 Total Beef, Pork, and Poultry Food Safety Articles, 1998 to 2005 57

Figure 5.1 ACF at Different Lag Lengths of all the Parameters in the Pooled SUR Tobit Model 151

Figure 5.2 ACF at Different Lag Lengths of all the Parameters in the Random Effects SUR Tobit Model 151

Figure A.1 Trace Plots of Media Index and Price Parameters from Pooled SUR Tobit Model 170

Figure A.2 Trace Plots of Media Index and Price Parameters from Random Effects SUR Tobit Model 171

Figure A.3 Histograms of Media Index and Price Parameters from Pooled SUR Tobit Model 172

Figure A.4 Histograms of Media Index and Price Parameters from Random Effects SUR Tobit Model 173

Figure A.5 Trace Plots of Price Parameters from Pooled SUR Tobit Model at Different Starting Values 174

Figure A.6 Trace Plots of Price Parameters from Random Effects SUR Tobit Model at Different Starting Values 175

1 Chapter

Overview and Objectives of Research

1.1 Overview

Many factors influence consumers' decisions to purchase meat and poultry products at the grocery store. One of those factors is food safety. It is estimated that bacterial pathogens cause approximately 5 million illnesses, 46,000 hospitalizations, and 1,458 deaths in the United States annually (Crump, Griffin, and Angulo, 2002). Pathogens including salmonella, E. coli, and listeria monocytogenes are naturally occurring in animals such as cattle, pigs, chickens, and turkeys. Consumer purchase decisions are also likely to be influenced by idiosyncratic experiences such as suffering from a foodborne illness or receiving medical warnings from a physician about their susceptibility to bacterial pathogens. However, general media information regarding 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. The reaction of consumers to changes in the amount of food safety information on beef, pork, and poultry available in the media is the focus of this study. Specifically, any differences in consumer reactions due to heterogeneous household characteristics are investigated.

Several studies have explicitly considered food safety effects on meat and poultry demand by employing various measures of media coverage to infer its effect on food demand (e.g. Burton and Young, 1996; Piggott and Marsh, 2004). These studies used aggregate data to jointly estimate meat and poultry demand equations that quantify the own- and cross-commodity effects of food safety information on purchases. This approach has shown that media information matters, but the effects are short-lived. The use of aggregate data assumes a representative consumer, so researchers cannot assess the likelihood or extent to which heterogeneous households might avoid purchasing meat and poultry products in response to food safety information. Examining this type of avoidance behavior at the disaggregate level will provide additional and complementary insight into the demand for food products under different levels of food safety information published in the media.

1.2 Objectives and Contributions

The objective of this study is to analyze the effects of food safety information across heterogeneous consumers by estimation of a demand model for meat and poultry that incorporates both food safety information and household characteristics. The question of whether or not differences in consumers affect their reactions to publicly available food safety information is investigated using both discrete and continuous models. The use of a discrete choice model allows for estimation of the likelihood that consumers will avoid purchasing meat or poultry products when the level of food safety information increases. A continuous demand model is also used to capture avoidance behavior by aggregating quantities purchased over a monthly time period and measuring marginal changes in consumption. This research question advances the existing food safety literature by explicitly considering consumer heterogeneity as it affects demand response to food safety information. Any differences in consumer response that could be attributed to heterogeneous household characteristics would be useful for understanding more fully what drives consumer response to general food safety information.

Another objective of this research is to employ an estimation strategy that accounts for the unique nature of household level purchase data including: censoring, panel data, and error correlation of demand system equations. Previous research has addressed these issues independently, but not in a single model. This study contributes to the existing literature by proposing a comprehensive estimation strategy that addresses each of these issues in the same model.

1.3 Outline of Dissertation

A review of the existing literature related to demand and food safety information is presented in Chapter 2. This review provides context for the current research and highlights how it contrasts with previous studies. An overview and brief literature review of studies estimating demand using household level data is also included in Chapter 2. This review outlines some of the more prominent issues that must be addressed when working with microeconomic data and how these issues have been handled in previous research.

Descriptions and summarizations of the data used in this study are provided in Chapter 3. The data source for the household purchase data is the Nielsen Homescan panel dataset. Access to this data is provided by the Economic Research Service of the U.S. Department of Agriculture through a cooperative research agreement. The data include information on meat consumption and expenditures as well as household-level demographic data. Additionally, this chapter includes a discussion on the prediction of missing prices due to censoring, the creation of a quality-adjusted price index, and quantity aggregation. Chapter 3 concludes with descriptions of the collection and creation of the food safety media index.

Chapter 4 includes the analysis of consumer demand in a discrete choice framework. Models of binary and multinomial choice situations are estimated. This analysis provides an intuitive basis for specification of the continuous demand model presented in Chapter 5. The continuous demand models are estimated using a seemingly unrelated system (SUR) tobit estimator. Explicit consideration of the panel nature of the data is made by estimation of both a pooled and random effects SUR tobit model. An overview of the study conclusions is presented in Chapter 6. This chapter also contains a discussion of the direction for future research, based on the results of this study.

2 Chapter

Literature Review

2.1 Demand and Food Safety Information

Several studies have focused on the effects of various types of food safety information and events on the demand for food commodities.[1] An early study of food safety impacts on consumer demand was conducted by Brown (1969) who looked at the effect of a health hazard “scare” from herbicide residue on cranberries. Information on the food safety event was considered to be a negative form of advertising. Brown argued that while positive advertising can make consumers less price responsive through increasing customer loyalty, negative information may cause them to become more price responsive. The adverse effects on cranberry demand were tested using comparisons of price elasticities of demand for the periods before, during, and after the event. No significant effect on price elasticity was found.

The use of media indices to measure the impact of food safety information on demand has been employed in several 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

[1] A brief summary of the research included in this literature review is provided in tables 2.1 and 2.2.

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 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. 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 analysis used quarterly data on quantity and expenditures for beef, lamb, pork, and poultry. 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. 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.

Marsh, Schroeder, and Mintert (2004) analyzed both media indices composed of newspaper articles and Food Safety and Inspection Service (FSIS) recall data as proxies for food safety information. Quarterly disappearance data from 1982 to 1998 on beef, pork, poultry, and other consumption goods was used to estimate an absolute price version of the Rotterdam model. Their findings indicated that while FSIS recall events significantly affect demand, media reports do not. However, the effect of recall events for beef and pork dies out quickly, within three periods, and effects are contemporaneous only for poultry recall events.

Food safety effects have also been considered outside the consumer demand model. Thomsen and McKenzie (2001) analyzed federally supervised meat and poultry recalls by publicly traded food companies from 1982 to 1998. They argue that, in addition to the costs of physically recalling meat, there may be adverse effects on stock values due to negative publicity or possible liability claims from food borne illnesses. The results of the study indicate that declines of 1.5 to 3 percent in shareholder values can occur from Class 1 food recalls, where there is a reasonable probability of severe health risks. Class 2 and 3 recalls, where the probability of a health risk is remote, did not have a significant effect on shareholder values. Similarly, Lusk and Schroeder (2000) considered the effect of meat recalls for beef and pork of various sizes and severity of health concern on nearby futures prices for live cattle and lean hogs, respectively. Their results found a marginally negative effect on short-term futures prices from medium-sized beef and large-sized pork Class 1 recalls.

Schlenker and Villas-Boas (2006) employed event study analysis to investigate the effects of media coverage of BSE on consumer and financial markets. Specifically, they considered the discovery of a BSE-positive cow in Washington State in December 2003 and health warnings about the potential effects of BSE aired on the Oprah Winfrey show seven years prior as the events of interest. 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. The CES, which tracks individual households for a two week period, did not reveal any statistically significant effect on consumer purchases or expenditures. The authors concluded that the repeated cross-section design of the CES may only provide an accurate measure of average purchasing decisions in a calendar year, while scanner and futures data may be more useful in tracking changes in buying habits over time.

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. Although some groups of targeted consumers responded to the advisory, there was little response from at-risk households that did not have high levels of education or newspaper readership. There was also found to be unintended spillover effects of decreased fish consumption among non-targeted households with high readership levels.

2.2 Demand Studies using Microeconomic Data

Estimation of meat demand using cross-sectional data allows for identification of demand determinants from household characteristics such as income, education, and age. However, panel data (time series and cross-sectional data) can be used to capture both the time aspect of demand decisions by consumers and the heterogeneous household effect. There are advantages to modeling the individual utility maximization decision using household level data, but several empirical problems can arise from the microeconomic data made available to researchers. Censored observations, price-quality variation, and quantity aggregation are all issues to address when using microeconomic data to estimate demand.

2.2.1 Censored Observations and Demand System Estimation

Household survey data includes both consuming and non-consuming households. The non-consuming households are censored observations of the consumer's decision to buy meat. These censored or zero observations make estimation of a demand system very challenging from an empirical perspective because it is difficult to find an estimator that accounts for censoring in a system framework. One approach to addressing this challenge is estimation of each demand equation separately, rather than as a system of equations (e.g., Cox and Wohlgenant, 1986; Capps and Park, 2002; Dong, Shonkwiler, and Capps, 1998). However, if cross-commodity effects exist, then a system of equations is required to estimate those joint effects efficiently. An alternative approach to single-equation estimation is to estimate the demand system by dropping observations on non-consuming households. However, useful information on the consumer's decision to buy meat would be lost and, depending on the level of censoring, a significant amount of the data would have to be omitted from analysis. Sample selection bias would be a concern in this situation because estimated effects from food safety may be over- or understated if the sample only includes people that purchase meat or poultry every period. There are several studies where microeconomic data was used that dealt with the estimation issues created by data censoring.

Perali and Chavas (2000) used cross-section data from Colombian urban households to estimate a censored AIDS model for several goods in two stages. The parameter restrictions required by theory for the demand model are set aside in the first stage and each commodity equation is individually estimated in unrestricted form. In the second stage, theoretically-consistent structural demand parameters are recovered by imposing cross-equation restrictions.

A similar technique was used by Meyerhoefer, Ranney, and Sahn (2005) to estimate a censored demand system using panel data on Romanian households. They first estimated the reduced form parameters and then used minimum distance estimation to identify the structural parameters and impose theoretical restrictions from the AIDS model. Their elasticity estimates differed from other studies using only cross-sectional data in that they were able to control for heterogeneous household preferences using random effects parameters that varied over time.

Golan, Perloff, and Shen (2001) proposed the use of Generalized Maximum Entropy as an alternative to Maximum Likelihood Estimation (MLE) for estimating censored demand systems. They estimated an AIDS model for five meat groups (beef, pork, chicken, processed meat, and fish) using data from a week-long survey of Mexican households. Dong, Gould, and Kaiser (2004) used the Amemiya-Tobin framework to estimate a censored AIDS model. They also used Mexican household data from a week-long survey period and estimated a system that included both meat and non-meat products. The authors employed a mapping rule from Wales and Woodland (1983) to ensure adding up of both the latent and observed expenditures shares in their demand system.

Double-hurdle models have also been used to estimate systems of censored demand equations (Heien and Wessells, 1990; Shonkwiler and Yen, 1999). The first step of a double-hurdle technique is estimation of a probit model to determine the probability of a purchase occurring. The second step is the estimation of a demand system that includes additional information from the first step in the form of an inverse Mills ratio. The two-step estimator has been popular in demand work because it is consistent under fairly strong assumptions about the error distribution and is easy to implement in applied work.

Gao and Spreen (1994) employed a double-hurdle model to estimate a hybrid Rotterdam demand system. Using household level budget data from 1987 to 88, they estimated a demand model for beef, pork, poultry, and fish that incorporated health information. The resulting empirical evidence supported the hypothesis that health concerns and convenience contributed to a structural change in meat demand away from beef and in favor of fish and poultry.

2.2.2 Price-Quality Variation

Prices pose a special challenge to researchers using microeconomic data. Most survey data only provides the researcher with quantity and expenditure information, so the price per unit of the commodity purchased must be imputed by dividing expenditures by quantity. Deaton (1989) referred to these as "unit values". The variation observed in these unit values may reflect more than supply shocks from transportation costs, cost of information, and seasonal variation. A portion of the variability in prices could be attributed to quality differences in the individual products that comprise the meat commodities measured by quantity demanded. Several methods have been proposed to account for price-quality variation in demand estimation.

Cox and Wohlgenant (1986) performed a demand analysis using cross-sectional household level budget data from the 1977-78 Nationwide Food Consumption Survey conducted by the U.S. Department of Agriculture (USDA). They noted that several previous studies using cross-sectional data assumed constant prices across consumers. However, significant variability in prices is observed in cross-sectional data and the constant prices assumption may lead to inaccurate demand estimates. Cox and Wohlgenant identified several sources of price variability including supply variability due to regional markets and quality differences resulting from heterogeneous commodity aggregation. To model the effect of quality on demand decisions, it was assumed that consumers first choose the quality of the commodity and then choose the quantity of the commodity they will buy. This allows the quality decision, as it affects price, to be modeled separately from the quantity demanded decision. The unit values were estimated as a function of regional/mean prices as well as household characteristics (income and family size) that may describe preferences for unobserved quality characteristics.

An alternative method of accounting for price-quality variation in household level data was put forth by Dong, Shonkwiler, and Capps (1998). They argued that both the price and quantity decisions are affected by sample selectivity in microeconomic data with censored observations. Therefore, rather than assuming separability of the price-quality and expenditure decisions, the two equations should be estimated simultaneously. Their methodology was supported by empirical evidence of a high level of correlation between the price-quality and the expenditure equations. However, this method used MLE of a joint density and may be computationally burdensome in some applications.

2.2.3 Quantity Aggregation

In addition to the problems created from quality variation in determining prices, problems can arise in quantity definitions. The study by Cox and Wohlgenant recognized the effect of quality variation on prices, but did not directly address the effect of quality on the definition of quantity demanded. They used a simple sum of physical quantities to comprise the composite good without a theoretical justification for the definition of quantity demanded. Nelson (1991) addressed this problem by noting that one of three different assumptions had to be made to justify simple sums of physical quantities to define quantity demanded.

The first assumption requires the elementary goods comprising the composite good to be perfectly substitutable. This makes quality differences irrelevant to the consumer's decision. This assumption only works for individual goods that are very homogeneous and is not a practical assumption for most analysis using cross-section data. The second alternative is the assumption of weakly separable preferences and homothetic with-in group preferences. This would imply that, at constant prices, the ratio of low to high quality items within a group would be the same for low and high income consumers (i.e., straight line income expansion paths). However, this assumption fails to justify the use of simple sums of physical quantities because it requires a quantity index that incorporates not only the quantity dimension, but all other aspects of the good that affect consumer valuation.

The third assumption, and the one Nelson suggests for problems using cross-section data, uses the Hicks composite commodity theorem. This theorem allows for the aggregation of individual goods consistent with freely variable choices across individual goods with varying characteristics. Under this assumption, aggregation of quantities is conducted by weighting the subgroups of the commodity aggregate by group average prices to account for the quality variation that is reflected in the prices.

2.3 Conclusion

A primary contribution of this proposed research will be the availability of household-level panel data. This type of data will allow for varying degrees of aggregation across time periods. Previous research has been limited by availability of quarterly, aggregate data making short-run effects from food safety information, if they exist, difficult to detect. It is also impossible to analyze reactions to food safety information as it varies across consumers using aggregate per capita consumption data. Therefore, heterogeneity of consumers can be incorporated into the demand analysis to provide further information on the effects of food safety information as it varies across different consumers.

Table 2.1 Summary of Demand and Food Safety Literature

Author	Research	Year and Publication	Measure of Information	Empirical Model	Key Findings
Demand Studies					
Brown	Effects on the demand for cranberries of a health scare from herbicide residue.	1969, American Journal of Agricultural Economics	Weekly household consumption data observed before, during, and after the food safety event.	Linear demand model	Food safety information did not have a statistically significant effect on demand.
Smith, van Ravenswaay, and Thompson	Effect on demand from media publicity of a heptachlor contamination of fresh fluid milk.	1988, American Journal of Agricultural Economics	Monthly per capita milk consumption data observed before and after the food safety event. Media index for comprised of newspaper articles rated as either negative or positive and dummy variables for event period.	Linear demand model with a polynomial lag structure for the media variables	Significant negative effects from negative news coverage. Positive news coverage did not significantly affect demand.
Dahlgran and Fairchild	Effect on demand from media publicity of salmonella contamination of chicken.	2002, Agribusiness	Weekly aggregate data on quantity and prices of chicken. Weekly newspaper articles and T.V. stories.	Linear ARIMA model that is non-linear in parameters	A significant negative demand response to adverse media was found, but it died out within a few weeks.
Burton and Young	Effect on demand for beef and other meats from BSE in Great Britain.	1996, Applied Economics	Quarterly data on quantities and expenditures. Media index included number of articles per quarter and the cumulative number of articles.	Dynamic Almost Ideal Demand System (AIDS)	Statistically significant short-run and long-run effects on the demand for beef and other meats were found.
Piggott and Marsh	Impact of food safety information on demand for beef, pork, and poultry.	2004, American Journal of Agricultural Economics	Quarterly disappearance data on U.S. per capita meat consumption and expenditures. Media indices are commodity specific.	Generalized Almost Ideal Demand System (GAIDS)	Impacts of food safety information were statistically significant, but did not last beyond the period in which the event occurred.
Marsh, Schroeder, and Mintert	Impact of food safety information and FSIS recalls on demand for beef, pork, poultry, and other consumption goods.	2004, Applied Economics	Quarterly disappearance and expenditure data on U.S. per capita beef, pork, poultry, and other consumption goods. FSIS recalls and media indices are commodity specific.	Absolute price version of Rotterdam model	FSIS recalls have a significant (although relatively short-lived) impact on demand, but media reports do not.

Table 2.1 Summary of Demand and Food Safety Literature, cont.

Event Studies					
Thomsen and McKenzie	Effects on stock values of publically traded companies from food recalls.	2001, American Journal of Agricultural Economics	Daily stock market returns were used to calculate shareholder values. FSIS recalls for meat and poultry, separated by recall class (severity).	Event study methods were used to compare actual returns to estimated returns	Class 1 recalls cause a 1.5 to 3 percent decline in shareholder values. Class 2 and 3 recalls do not have a statistically significant effect.
Lusk and Schroeder	Effects on futures prices for cattle and hogs from food recalls of beef and pork.	2000, NCR-134 Conference Proceedings	Daily futures prices for live cattle and lean hog futures contracts. FSIS recalls for meat and pork, separated by class and the amount of product recalled.	Event study methods were used to compare actual to estimated futures contract prices	Medium-sized beef and large-sized pork Class 1 recalls have a negative, short-term impact on futures prices.
Schlenker and Villas-Boas	Responses of consumer and financial markets to media coverage of BSE.	2006, University of California-Berkeley, CUDARE Working Paper	UPC scanner data and diary files from the U.S. Consumer Expenditure Survey (CES) data of daily purchases of meats. Daily futures prices for live cattle.	Event study methods were used to compare effects across the different data sources.	Statistically significant and negative effects were found using the scanner and futures data. No effect was found using the CES data.
Shimshack, Ward, and Beatty	Consumer responses to FDA advisories on exposure to methyl-mercury from store-bought fish.	2007, Journal of Environmental Economics and Management	Household level data was used from the U.S. Consumer Expenditure Survey.	Reduced form analysis employing both parametric and non-parametric methods.	Short-run responses to the FDA warnings existed and were primarily determined by education and newspaper readership. Spillover effects to non-targeted groups also occurred.

Table 2.2 Summary of Demand Research using Microeconomic Data

Author	Research	Year and Publication	Measure of Information	Empirical Model
Gao and Spreen	Effects of health information on the demand for beef, pork, poultry, and fish	1994, Canadian Journal of Agricultural Economics	Household level consumption data collected over a one week period.	Hybrid demand system combining the generalized addilog system (GADS) and level version of Rotterdam model
Perali and Chavas	A demand system is estimated for food, housing, health, education, clothing, transportation, and all other goods.	2000, American Journal of Agricultural Economics	Household level consumption data collected over a one week period.	Tobit model using an AIDS specification
Meyerhoefer, Ranney, and Sahn	A demand system is estimated for various food groups, gasoline, and non-foods.	2005, American Journal of Agricultural Economics	Three years of household level consumption data collected over a one month period.	AIDS model with a random effects specification
Golan, Perloff, and Shen	A demand system is estimated for five meat groups.	2001, Review of Economics and Statistics	Household level consumption data collected over a one week period.	AIDS model
Dong, Gould, and Kaiser	A demand system is estimated that includes both meat and non meat products.	2004, American Journal of Agricultural Economics	Four months of weekly household level consumption data.	Tobit model using an AIDS specification

3 Chapter

Meat and Poultry Consumption and Food Safety Data

3.1 Introduction

This chapter describes the data used to analyze the effects of food safety information on U.S. household demand for meat and poultry. Monthly data are used for this analysis from the time period January 1998 to December 2005. In this chapter, details are given on the creation of the variables used in subsequent chapters.

The data for this study come from three 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.

3.2 U.S. Household Consumption Data

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 each 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.

3.2.1 Meat and Poultry Purchases

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.[2] 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.

[2] Examples of processed meat and poultry products include luncheon meats, frozen dinners, or soups that contain meat or poultry.

Initial inspection of the daily transaction data indicated possible outlier observations or reporting errors. Therefore, prior to finalizing the dataset, several rules were developed to eliminate these problematic outlier observations. All duplicate purchases were deleted from the dataset. This was done based on visual inspection of the data that suggested these purchases were incorrectly recorded. Observations where the total quantity purchased was less than 0.25 pounds were also deleted from the dataset. This rule was used because all the purchases that met this criterion appeared to be reporting errors. The data also contained some extremely high per unit prices. These may be due to recording errors or possibly highly specialized meat purchases (e.g. mail order or home delivery). In order to determine a reasonable rule for deleting high prices, the individual products within each commodity group were analyzed to determine their respective price distributions. For each commodity, the upper one percent of the distribution of the highest priced product was used as a cut off value. This cut off price is $36.45/lb for beef, $18.14/lb for pork, and $20.64/lb for poultry.[3] Very low prices are also present in the dataset, due to the use of coupons for some purchases. Coupon value ranges from zero up to 100 percent of the total price of a product, making some prices equal to zero. However, because these are valid purchases, they were not removed from the dataset.[4] After these data cleaning rules were implemented, approximately 1.85% of the beef purchases, 5.80% of the pork purchases, and 1.76% of the poultry purchases were discarded from the dataset. The final sample for beef and veal consists of

[3] As a comparison, the commodity average prices are $3.01/lb for beef, $2.48/lb for pork, and $1.95/lb for poultry.

[4] The percentage of all beef, pork, and poultry purchases where a coupon accounts for the full price is approximately 0.13%.

1,321,058 observations. Pork purchases total 487,748 observations over the entire sample period and the final sample of poultry purchases is comprised of 811,840 observations. Combined, there are 2,620,646 purchase transactions for beef, pork, and poultry in the final sample.

Summary statistics of household level purchases, expenditures, and retail prices for beef, pork, and poultry are presented in table 3.1. The purchase data were grouped into five beef, four pork, and six poultry products having similar characteristics and average prices. While some quality and price variation still exists within these groups (e.g. all grades of steak are included in the beef steak category), the level of variation is much smaller than it would be if purchases were aggregated into groups of beef, pork, and poultry.

The product groupings for beef are ground beef, roasts, steaks, frozen, and other.[5] Ground beef make up over 41% of the total beef purchases and have the lowest average retail price of $2.14 per pound. Steaks have an average price of $4.56 per pound, which is the highest price of all the beef, pork, and poultry products. Pork chops are the largest group of pork purchases (53.2%) and also have the highest average price of all pork products at $2.93 per pound. Ground pork has the lowest price per pound at $2.01 and also comprises the smallest percentage of all pork purchases. Whole chicken and bone-in chicken pieces are the most commonly purchased poultry group, making up 44.6% of poultry purchases. This group's average retail price is $1.29 per pound. Turkey is the only product group with a lower price per pound of $1.26. The highest priced poultry group is boneless chicken and

---

[5] The frozen products do not include any further processing beyond the freezing process. The 'other' category includes fresh products that do not fall into the other categories, such as ribs, stew meat, liver, etc.

turkey pieces. The average price of this group is $2.69 per pound. With the exception of frozen chicken and turkey, which make up over 10% of all poultry purchases, frozen meat products are a very small proportion of all meat and poultry purchases. The frozen beef category makes up the smallest percentage of all beef purchases and frozen pork purchases were too small to be considered a separate category.[6]

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 are aggregated to a weekly level, the amount of censoring in this dataset is very large.[7] 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 where removed from the panel, leaving 62,136 households across all eight sample years. Although households that made no purchases of meat and poultry were removed, a large amount of censoring remains in the monthly data. The percentages of censored observations for monthly purchases of beef, pork, and poultry are presented in table 3.2. Beef products have the lowest amount of censoring across all the sample years. Of the 745,632 monthly household observations, 42.2% have a zero quantity. This means that 57.8% of households bought at least one beef product on a monthly basis.

[6] Frozen pork purchases comprise less than 1% of all pork purchases and were grouped into the other pork category.
[7] The percentages of censored observations when data are aggregated using a weekly periodicity are 77.5% for beef purchases, 89.4% for pork purchases, and 84.2% for poultry purchases.

Poultry purchases are censored an average of 52.27%, followed by pork purchases at 65.12%.

The percentages of all monthly household observations that are censored, by product group, are presented in table 3.3. The product groups that are bought most often on a monthly basis by households in the sample are ground beef, pork chops, whole chicken, and bone-in chicken pieces. The product groups that are bought with the least amount of frequency are ground pork, frozen beef, and ground poultry.

3.2.2 Household Demographics

The Nielsen Homescan panel is a stratified random sample that was selected based on both geographic and demographic targets. Participation rates of households in the sample are listed for each year in table 3.4. Annual participation in the panel ranged from a low of 6,966 households in 1999 to a high of 8,428 households in 2003. Participation across sample years ranged from one to eight years, with the largest percentage of households participating for one year of the panel (32.7%). To be considered a participant for a sample year, the household must have participated for at least 10 of 12 months of the year.

The dataset used in this study is an unbalanced panel 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. The values and frequencies of variables describing these household characteristics are listed in table 3.5 and were calculated by averaging across all sample years. The characteristics are also summarized graphically in figure 3.1, where the percentage of the total sample is given for each demographic category. Two person households comprise just over 37.5% of the total sample, while households with five or more members make up only 9.4% of the sample. Most households in the sample (70.3%) do not have children under the age of 18 living in the house. Income level appears to be relatively evenly distributed across the mid-range income levels ($20,000 to $49,999 per year), with a large percentage falling in the range of $70,000 to $99,999 per year. The age distribution indicates few participants under age 30, while the largest percentage of household heads are in the 55 to 64 year age range. Most head of household in the sample are employed full time (over 35 hours per week). However, a relatively large percentage of heads of household are not employed for pay (18.8 % of men and 32.7% of women). The highest education level attained by the head of household is relatively evenly distributed between high school, some college, and college graduates. Over 61% of the sample is comprised of married households, which allows for information on both the male and female head of household. The sample is relatively evenly distributed over the four geographical regions, with the highest participation in the southern region.

3.2.3 Demographics and Purchases of Select Groups

In addition to analyzing the data across the entire sample, it is useful to summarize the characteristics of select groups based on certain characteristics. The first of the three groups analyzed in this section is high income households, which reported an annual household income of $100,000 or more. This group is made up of 5,642 households over the eight sample years. The second group is comprised of households with children under the age of 18 living at home. This group consists of 18,364 households over the entire sample period. The third subgroup that was analyzed is households with both heads aged 55 years and older. There are 12,467 observations over the sample period for this group. These groups are not mutually exclusive. That is, a household with children may be included in the high income group, a high income household may have both heads aged 55 and older, and so on.

The Nielsen dataset offers a wide array of options for analyzing the data using demographic characteristics. The three groups used for this analysis were selected due to their unique characteristics with respect to food safety information and the demand for meat and poultry. High income households may face a larger number of affordable substitutes, which could affect their consumption patterns in the presence of food safety information. These households also tend to have higher education levels, which could affect the way in which they process and respond to food safety information. Households with children and people aged 55 and older are interesting groups due to their potential risk level with respect to food borne pathogens. The two groups are the most susceptible to becoming very sick or even dying from exposure to these pathogens. Therefore, the demand response of these households to food safety information may be different than households without these relatively higher risk members.

Household participation rates by group are presented in table 3.6. Participation across sample years is relatively constant and similar to that of the whole sample. The high income households and those with children present have across-year participation patterns like those of the whole sample. Approximately 45% of the households in these groups participated for one year, while only about 3% participate for all eight years. Households with heads aged 55 and older have a slightly different participation pattern. The majority of these households only participate for one year (25.6%). However, there is a much larger percentage that participated for several years as compared to the other groups, with over 10% participating for all eight years.

The demographic and geographic characteristics of the three subgroups of households are summarized in table 3.7. A graphical summary of the characteristics is also presented in figure 3.2, where the percentage of households within subgroups is given for each demographic category. There are a few distinct differences in household characteristics between the three groups. First, households with heads aged 55 and older make up the largest percentage of both the male and female heads not employed for pay. This is expected, given that the age range includes people likely to be retired. They are also the group with the largest percentage of education levels below college graduate. High income households are comprised predominately of four or less people and make up the highest percentage of college and post college graduates for both male and female heads. High income households also have the highest percentage of both men and women working full time, as compared to the other groups. Households with children present have the largest percentage of female heads of household working less than 35 hours per week. This is likely due to mothers staying at home to care for children either full or part time.

Other interesting comparisons that can be made between these three groups are their meat and poultry purchases. The monthly average purchases made by these groups across all sample years are presented in table 3.8. These monthly averages give some idea of the relative quantities of various meat and poultry products that these groups purchase. Relative to the other groups, households with children present purchase larger quantities of ground beef, whole and bone-in chicken pieces, and frozen chicken and turkey. Given the time and income constraints that are common for families with children, it makes sense that they would purchase products which tend to be lower priced and more convenient to prepare than other meat and poultry products. Households with heads aged 55 and older purchase more beef roasts, pork roasts and hams, and whole and bone-in turkey, relative to the other groups. Given the high percentage of these households with heads that are not employed for pay, their opportunity cost of time for food preparation may be lower than the other groups. Therefore, they tend to purchase products that require more preparation and cooking time.

High income households, on average, purchase slightly more ground and boneless poultry and less beef roasts, pork chops, and whole chicken than the other groups. Both ground and boneless poultry are relatively high priced products. These households do not appear to purchase larger amounts of other high prices products such as beef steak. However, this dataset includes only purchases made at grocery stores and similar food stores. If higher income households are eating more of their meals away from home than the other groups, then their meat and poultry purchases from grocery stores may not fully reflect the income effect that would be expected for certain products with higher average prices.

3.3 Missing Prices

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.

Several studies have used regional average prices paid by consuming households to replacing missing prices for non-consuming households, often with some kind of adjustment for degree of urbanization or household-specific preferences (Cox and Wohlgenant, 1986; Dong, Gould, and Kaiser, 2004; Golan, Perloff, and Shen, 2001). In this study, a similar approach is employed that uses sample averages of monthly prices paid by consuming households for each beef, pork, and poultry group, as well as regional and demographic characteristics of the households to impute missing prices.[8] Following Cox and Wohlgenant, 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

[8] The use of an ad-hoc method of filling in missing prices may lead to selection bias in price elasticities of demand. Alternative estimation strategies for dealing with selection bias include the Heckman’s (1979) two-step estimator or the method of joint estimation of purchase choice and prices put forth by Erdem, Keane, and Sun (1999). However, the choice of the most appropriate method for this study is an empirical one and left for future research.

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:

$$p_{itn} = \alpha \overline{p}_{it} + \boldsymbol{\gamma}_r \mathbf{r}_n + \delta u_n + \eta i_n + \kappa i_n^2 + \tau s_n + \rho s_n^2 + \varepsilon_{it} \ , \tag{3.1}$$

where p_{itn} is the observed price of good *i* in month *t* for consuming household *n*, $\overline{p}_{it}$ 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, ε_{it} is an iid error term, and $\alpha, \boldsymbol{\gamma}_r, \delta, \eta, \kappa, \tau,$ and ρ are the corresponding coefficients to be estimated.[9] 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 results of the price regressions for each good are presented in table 3.9. The coefficients of most of the variables are statistically significantly different from zero at the 5 percent level across all the price regressions. All the coefficients for the monthly average group price are positive, while the signs of the regional binary variables vary, depending on

[9] Total household income is recorded as an interval in this dataset (see table 3.5). 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. This method of converting intervals to continuous values may result in inconsistent estimates (Stewart, 1983) and the degree of bias increases with the degree of the interval (Cameron, 1987). However, the income intervals given in this dataset have a relatively small range, so the effect of the midpoint method on parameter consistency is expected to be small.

the regression. The coefficient for urbanization, u_n, is positive for all the goods, except frozen beef and poultry. The positive sign indicates that consumers in urban areas tend to pay slightly more for meat and poultry products than consumers in non-urban areas. The coefficients for household income, i_n, are positive across all equations and the coefficients for i_n^2 are negative for all but one good. The negative sign on the quadratic income term indicates that, while higher income households pay more for meat and poultry products, the effect declines as income rises over the relevant range of income values. The coefficients for household size were the expected negative sign and the squared terms tended to be positive when they were statistically significantly different from zero. These results suggest that size economies for purchases of meat and poultry have a downward effect on prices that diminish as household size increases.

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 summary statistics of monthly group prices are displayed in table 3.10 for both the observed and predicted prices of each good.

3.4 Price Indices

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.

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 price index 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 price index is specified following Törnqvist (1936) and is an expenditure share-weighted geometric price index defined as:

$$p_{nt}^{B} = \prod_{i=1}^{G} p_{\mathrm{int}}{}^{w_i} \quad , \tag{3.2}$$

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:

$$w_i = \frac{\bar{p}_i \bar{x}_i}{\sum_{j=1}^{G} \bar{p}_j \bar{x}_j} \ , \tag{3.3}$$

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.[10] 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. Summary statistics of the beef, pork, and poultry geometric price indices are presented in table 3.11.

3.5 Quality Adjusted Quantities

Aggregation of the total quantities purchased to a commodity level may diminish the quality variation that exists between the individual meat and poultry products. Therefore, some adjustment to the aggregation process for quantities that accounts for quality variation is needed. Following Nelson (1991), the Hicks composite commodity theorem is assumed and average prices of the individual meat and poultry groups are used to represent quality differences. These group average prices can be used to weight the individual group quantities within the aggregation process. The quality adjusted quantities are defined as follows:

$$q_{nt}^{B} = \sum_{i=1}^{G} q_{nt}^{g_i} * \bar{p}^{g_i} \ , \tag{3.4}$$

[10] 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.

where q_{nt}^{B} is the quality adjusted quantity of beef purchased by household *n* in month *t*, $q_{nt}^{g_i}$ is the monthly quantity purchased of beef group *i* by the household, $\overline{p}^{g_i}$ is the average price of beef group *i* across the entire sample period, and *G* is the number of subgroups specified for beef. A similar price index was also calculated for the pork and poultry quantity aggregates. Prior to estimation, the quality adjusted quantities are divided by the number of people in the household to arrive at a per capita quantity purchased measure. No adjustment could be made for the number of adults versus children because sufficient information to such an adjustment was not available. Summary statistics of the quality adjusted per capita quantities are presented in table 3.11 for beef, pork, and poultry.

3.6 Food Safety Media Indices

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 names and locations of the regional newspapers are given in table 3.12. 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*.

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. The intuition for this specification of the indices is that each day of the month 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 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.[11]

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. Summary statistics of household purchase frequency are given in table 3.13 and a histogram of the purchase frequencies is provided in figure 3.3. The bi-weekly frequency with which households purchase meat and poultry serves as the basis for a two week memory specification of the media index. Summary statistics for each commodity index, by region, are shown in table 3.14 for the years 1998 to 2005. Figures 3.4, 3.5, and 3.6 display the regional media indices for beef, pork, and poultry, respectively. While the indices tend to follow the same trends overall, there is some variation in the indices that reflects differences in regional media coverage of food safety information.

[11] Initial model estimation was conducted with a food safety media index that was a simple linear aggregation to the monthly level, as used in Piggott and Marsh (2004). The parameter estimates of food safety information in these preliminary models did not change appreciably depending on the specification of the media index.

The total number of newspaper articles for beef, pork, and poultry is displayed in figure 3.7. The level of food safety articles is relatively constant during most months, with noticeable spikes in articles for beef in March 2001, December 2003, and January 2004. The large number of articles in March 2001 corresponds to an outbreak of foot and mouth disease in Europe, while the large number of articles in December 2003 and January 2004 are a result of the discovery of a BSE-positive cow in Washington State. The large spikes in the poultry index in January through February of 2004 and October through November of 2005 correspond to outbreaks of avian influenza in several Asian countries as well as reports of poultry to human infection that was often fatal. There was also a large amount of news articles covering the U. S.'s policy for dealing with avian influenza if found in domestic flocks. The index for pork is made up of much fewer articles, but still displays some spikes in news coverage. These periods of increased food safety are usually correlated with beef or poultry events. [12] Despite the absence of a large food safety event specific to pork, the sample period did contain instances of pork products being subject to recalls due to listeria and salmonella.

3.7 Conclusion

The use of data on household level meat and poultry purchases and species-specific media indices allows for the investigation of the demand effects of food safety information as it varies across consumers. This dataset contains monthly data from 1998 to 2005. A monthly

[12] The correlation between the indices is as follows: beef-pork (0.613), beef-poultry (0.517), and pork-poultry (0.248).

periodicity allows for analysis of relatively short-run demand effects from food safety information as well as mitigating the number of censored observations common in household level data. The data were sorted into fifteen sub-species groups having similar product characteristics and average prices. The grouping methodology is intended to reduce quality variation that would be found in using simple beef, pork, and poultry aggregates.

Further work with the data included estimation of the missing retail prices for households that did not make purchases during a given month. Monthly average prices for each subgroup, regional effects, and income and household size effects were used to estimate these prices. The observed and estimated prices were subsequently used to create monthly price indices for beef, pork, and poultry aggregate commodities.

Initial investigation of the demographics and consumption patterns of three subgroups of the Nielsen panel indicate that observable household heterogeneity plays an important role in demand estimation. The three subgroups (high income households, households with children present, and head of households aged 55 and older) have distinct differences in the product groups of meat and poultry they consume. Households with heads aged 55 and older tend to purchase products requiring longer preparation times (e.g. beef roast and hams), while households with children tend toward lower priced and convenient preparation meat and poultry products (e.g. ground beef and chicken pieces). Variation between the subgroups in characteristics such as education and employment status also suggests possible differences in their responses to food safety information. Therefore, the reduced-form demand analysis presented in the next chapter will focus not only on the full sample of households, but also certain subgroups to determine if demand response to food safety information varies across heterogeneous consumers.

Table 3.1 Summary Statistics of Household Meat and Poultry Purchases

	Percent of Group[a]	Average	Minimum[b]	Maximum	Std. Dev.
BEEF					
Ground beef	41.11				
Quantity (lbs)		2.211	0.250	100.000	1.634
Expenditure ($)		4.265	0.000	104.800	2.988
Retail Price ($/lb)		2.141	0.000	36.361	0.980
Roasts	13.82				
Quantity (lbs)		3.126	0.250	68.600	1.865
Expenditure ($)		7.260	0.000	103.920	5.797
Retail Price ($/lb)		2.504	0.000	36.362	1.464
Steaks	30.55				
Quantity (lbs)		1.660	0.250	96.800	1.404
Expenditure ($)		6.730	0.000	216.500	6.248
Retail Price ($/lb)		4.556	0.000	36.355	2.980
Frozen	1.32				
Quantity (lbs)		2.448	0.313	64.000	1.979
Expenditure ($)		6.741	0.000	97.160	4.540
Retail Price ($/lb)		3.148	0.000	19.980	1.396
Other	13.21				
Quantity (lbs)		2.113	0.250	349.000	2.149
Expenditure ($)		4.770	0.000	611.480	4.572
Retail Price ($/lb)		2.661	0.000	36.393	1.860
PORK					
Ground pork	2.52				
Quantity (lbs)		1.391	0.250	32.000	1.187
Expenditure ($)		2.589	0.000	37.400	1.882
Retail Price ($/lb)		2.013	0.000	17.447	0.933
Roasts and hams	20.28				
Quantity (lbs)		4.861	0.250	165.900	3.478
Expenditure ($)		9.120	0.000	220.430	6.532
Retail Price ($/lb)		2.448	0.000	18.136	1.848
Steaks and chops	53.24				
Quantity (lbs)		2.219	0.250	119.860	2.017
Expenditure ($)		5.294	0.000	94.170	3.479
Retail Price ($/lb)		2.934	0.000	18.030	1.515
Other	23.95				
Quantity (lbs)		2.269	0.250	120.000	2.831
Expenditure ($)		4.363	0.000	99.990	3.848
Retail Price ($/lb)		2.417	0.000	18.137	1.726

[a] Percent of group calculated as a percentage of the beef (1,321,058), pork (487,748), and poultry (811,840) group purchases, respectively.

[b] Zero expenditure and price values are due to the use of coupons.

Table 3.1 Summary Statistics of Household Meat and Poultry Purchases, cont.

	Percent of Group[a]	Average	Minimum[b]	Maximum	Std. Dev.
POULTRY					
Ground poultry	4.29				
Quantity (lbs)		1.697	0.260	30.000	1.341
Expenditure ($)		3.413	0.000	82.130	2.337
Retail Price ($/lb)		2.190	0.000	20.000	1.116
Whole, bone-in chicken	44.58				
Quantity (lbs)		4.021	0.250	695.250	2.933
Expenditure ($)		4.155	0.000	300.000	2.696
Retail Price ($/lb)		1.291	0.000	20.333	1.030
Whole, bone-in turkey	7.82				
Quantity (lbs)		8.032	0.250	171.510	6.204
Expenditure ($)		7.543	0.000	81.260	5.752
Retail Price ($/lb)		1.262	0.000	20.423	1.090
Boneless	25.92				
Quantity (lbs)		2.505	0.250	73.600	1.727
Expenditure ($)		5.641	0.000	99.900	3.648
Retail Price ($/lb)		2.694	0.000	20.505	1.451
Frozen	10.23				
Quantity (lbs)		4.280	0.250	80.000	3.075
Expenditure ($)		7.853	0.000	111.920	5.169
Retail Price ($/lb)		1.959	0.000	17.320	0.904
Other	7.16				
Quantity (lbs)		2.213	0.250	824.560	4.359
Expenditure ($)		3.584	0.000	190.000	3.345
Retail Price ($/lb)		2.204	0.000	20.636	1.891

[a] Percent of group calculated as a percentage of the beef (1,321,058), pork (487,748), and poultry (811,840) group purchases, respectively.

[b] Zero expenditure and price values are due to the use of coupons.

Table 3.2 Censored Monthly Observations by Year

Year	Beef	Pork	Poultry
1998	36.84	60.54	48.94
1999	37.40	61.11	48.81
2000	39.21	64.49	49.91
2001	42.23	65.31	51.68
2002	43.80	66.34	53.00
2003	45.42	66.79	54.02
2004	46.36	67.35	55.20
2005	46.32	69.01	56.59

Table 3.3 Censored Monthly Observations by Commodity[a]

Group	Average
BEEF	
Ground beef	62.46
Roasts	82.89
Steaks	72.73
Frozen	97.94
Other	85.46
PORK	
Ground	98.68
Roasts and hams	89.28
Steaks and chops	77.63
Other	89.89
POULTRY	
Ground poultry	96.82
Whole, bone-in chicken	74.54
Whole, bone-in turkey	93.42
Boneless	83.20
Frozen	91.25
Other	94.50

[a] Summary statistics are across the entire sample from 1998 to 2005.

Table 3.4 Household Participation Rates

Within-Year Participation			Across-Year Participation		
Year	**Number of Households**	**Percent**	**Years in Panel[a]**	**Number of Households**	**Percent**
1998	7,465	12.01	1	6,159	32.74
1999	6,966	11.21	2	3,200	17.01
2000	7,307	11.76	3	2,233	11.87
2001	7,890	12.70	4	1,664	8.85
2002	8,284	13.33	5	1,525	8.11
2003	8,428	13.56	6	1,245	6.62
2004	8,043	12.94	7	1,161	6.17
2005	7,753	12.48	8	1,625	8.64

[a] Households participated for at least 10 months of the sample year.

Table 3.5 Household Panel Demographic Variables

Demographic Variable	Frequency	Percent of Sample[a]
Household Size		
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
Household Income		
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
Age of Male Head[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
Age of Female Head[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

[a] Summary statistics calculated as average across the eight sample years.

[b] Married households have information on both the male and female head of household.

Table 3.5 Household Panel Demographic Variables, cont.

Demographic Variable	Frequency	Percent of Sample[a]
Age and Presence of Children		
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
Male Head Employment[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
Female Head Employment[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
Male Head Education[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
Female Head Education[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
Region		
East	1,658	21.32
Central	1,582	20.53
South	2,840	36.45
West	1,687	21.70
Marital Status		
Married	4,755	61.37
Widowed	618	7.90
Divorced/Separated	1,142	14.64
Single	1,253	16.09

[a] Summary statistics calculated as average across the eight sample years.

[b] Married households have information on both the male and female head of household.

Table 3.6 Household Participation Rates by Demographic Group

Within-Year Participation			Across-Year Participation		
Year	Number of Households	Percent	Years in Panel[a]	Number of Households	Percent
High Income					
1998	6,324	9.34	1	1,029	44.49
1999	5,976	8.83	2	504	21.79
2000	6,852	10.12	3	266	11.50
2001	8,136	12.02	4	170	7.35
2002	9,696	14.32	5	130	5.62
2003	9,648	14.25	6	93	4.02
2004	10,296	15.21	7	49	2.12
2005	10,776	15.92	8	72	3.11
Children Present					
1998	32,916	14.94	1	3,430	45.45
1999	27,168	12.33	2	1,509	19.99
2000	26,856	12.19	3	921	12.20
2001	29,088	13.20	4	556	7.37
2002	29,832	13.54	5	465	6.16
2003	27,732	12.58	6	271	3.59
2004	24,324	11.04	7	182	2.41
2005	22,452	10.19	8	213	2.82
55 Years and Older					
1998	14,748	9.86	1	883	25.63
1999	15,396	10.29	2	602	17.47
2000	16,524	11.05	3	468	13.58
2001	18,108	12.10	4	342	9.93
2002	19,524	13.05	5	290	8.42
2003	21,360	14.28	6	232	6.73
2004	21,600	14.44	7	258	7.49
2005	22,344	14.94	8	370	10.74

[a] Households participated for at least 10 months of the sample year.

Table 3.7 Household Panel Demographic Variables by Demographic Group

Demographic Variable	High Income	Children Present	55 and Older
		Percent of Sample [a]	
Household Size			
Single member	7.55	0.00	0.00
Two members	41.19	5.50	79.44
Three members	19.80	27.17	13.52
Four members	19.74	38.34	4.53
Five members	7.99	18.68	1.45
Six members	2.61	6.55	0.72
Seven members	0.73	2.43	0.22
Eight members	0.25	0.76	0.09
Nine or more members	0.14	0.57	0.03
Household Income			
Under $5000	--	0.41	0.24
$5000-$7999	--	0.54	0.20
$8000-$9999	--	0.47	0.28
$10,000-$11,999	--	0.88	0.82
$12,000-$14,999	--	1.66	1.72
$15,000-$19,999	--	3.44	4.60
$20,000-$24,999	--	5.45	7.22
$25,000-$29,999	--	5.45	7.46
$30,000-$34,999	--	6.48	8.51
$35,000-$39,999	--	6.44	8.07
$40,000-$44,999	--	7.61	8.15
$45,000-$49,999	--	6.88	7.58
$50,000-$59,999	--	12.92	12.03
$60,000-$69,999	--	11.73	9.80
$70,000-$99,999	--	19.05	15.12
$100,000 & Over	100.00	10.59	8.19
Age of Male Head [b]			
Under 25 Years	0.02	0.43	--
25-29 Years	0.80	3.34	--
30-34 Years	6.15	11.58	--
35-39 Years	10.39	17.74	--
40-44 Years	13.35	19.72	--
45-49 Years	15.19	16.22	--
50-54 Years	19.00	8.85	--
55-64 Years	21.68	6.02	42.89
65+ Years	7.48	1.70	57.11
No Male Head	5.96	14.39	0.00
Age of Female Head [b]			
Under 25 Years	0.14	1.14	--
25-29 Years	1.77	6.00	--
30-34 Years	7.64	15.96	--
35-39 Years	11.70	22.69	--
40-44 Years	14.25	22.60	--
45-49 Years	16.91	16.05	--
50-54 Years	17.25	7.59	--
55-64 Years	17.69	4.66	58.11
65+ Years	5.16	1.21	41.89
No Female Head	7.50	2.10	0.00

[a] Summary statistics calculated as average across the eight sample years. Sample size is 5,642 observations for high income households, 18,364 observations for households with children, and 12,467 oberservations for head of households aged 55 years and older.

[b] Married households have information on both the male and female head of household.

Table 3.7 Household Panel Demographic Variables by Demographic Group, cont.

Demographic Variable	High Income	Children Present	55 and Older
	Percent of Sample [a]		
Age and Presence of Children			
Under 6 only	5.41	14.39	0.77
6-12 only	7.27	23.89	1.11
13-17 only	10.60	27.37	1.90
Under 6 & 6-12	4.71	13.15	0.27
Under 6 & 13-17	0.51	2.07	0.07
6-12 & 13-17	5.26	16.19	0.29
Under 6 & 6-12 & 13-17	0.69	2.94	0.03
No Children Under 18	65.54	--	95.55
Male Head Employment [b]			
Under 30 hours	2.45	1.69	6.54
30-34 hours	2.00	1.79	2.82
35+ hours	79.67	75.90	29.54
Not Employed for Pay	9.93	6.23	61.10
No Male Head	5.96	14.39	0.00
Female Head Employment [b]			
Under 30 hours	10.44	17.11	11.37
30-34 hours	3.86	6.56	4.20
35+ hours	56.75	45.59	21.07
Not Employed for Pay	21.45	28.63	63.36
No Female Head	7.50	2.10	0.00
Male Head Education [b]			
Grade School	0.05	0.87	2.77
Some High School	0.67	4.12	7.05
Graduated High School	6.35	20.64	24.21
Some College	19.21	26.27	29.20
Graduated College	34.37	23.77	21.62
Post College Grad	33.39	9.93	15.16
No Male Head	5.96	14.39	0.00
Female Head Education [b]			
Grade School	0.05	0.41	0.91
Some High School	0.32	2.39	4.53
Graduated High School	7.82	23.25	31.85
Some College	22.97	33.03	35.24
Graduated College	36.48	30.09	18.35
Post College Grad	24.87	8.75	9.11
No Female Head	7.50	2.10	0.00
Region			
East	23.79	21.96	19.61
Central	14.85	20.21	21.19
South	35.75	37.38	36.07
West	25.61	20.45	23.13
Marital Status			
Married	82.88	80.40	97.08
Widowed	2.11	2.16	1.06
Divorced/Separated	5.64	11.16	1.15
Single	9.38	6.28	0.71

[a] Summary statistics calculated as average across the eight sample years. Sample size is 5,642 observations for high income households, 18,364 observations for households with children, and 12,467 oberservations for head of households aged 55 years and older.

[b] Married households have information on both the male and female head of household.

Table 3.8 Monthly Meat and Poultry Purchases by Demographic Group

	High Income	Children Present	55 Years and Older
	Average Monthly Purchase (lbs)		
BEEF			
Ground beef	1.40	2.14	1.79
Roasts	0.80	0.86	1.06
Steaks	1.04	1.02	1.04
Other	0.43	0.59	0.57
Frozen	0.05	0.09	0.05
PORK			
Ground pork	0.03	0.02	0.03
Roasts and hams	0.64	0.65	1.02
Steaks and chops	0.76	0.91	0.94
Other	0.25	0.38	0.45
POULTRY			
Ground poultry	0.11	0.09	0.08
Whole, bone-in chicken	1.88	2.39	2.12
Whole, bone-in turkey	0.83	0.70	0.98
Boneless	1.00	0.94	0.64
Other	0.16	0.19	0.19
Frozen	0.43	0.71	0.43

Table 3.9 Estimated Coefficients of Censored Product Price Models

	Monthly Average Group Price	Region East	Central	South	West	Urban Area	Household Income	Household Income^2	Household Size	Household Size^2
BEEF										
Ground beef										
Coefficient	0.823*	0.288*	0.083*	0.159*	0.369*	0.132*	7.6E-06*	-2.2E-11*	-0.120*	0.009*
Robust Std. Error	(0.006)	(0.017)	(0.017)	(0.017)	(0.017)	(0.005)	(2.1E-07)	(1.6E-04)	(0.005)	(0.001)
R^2	0.860									
Roasts										
Coefficient	0.942*	-0.044	-0.180*	-0.059	0.020	0.101*	6.2E-06*	-4.2E-12	-0.078*	0.005*
Robust Std. Error	(0.012)	(0.038)	(0.037)	(0.038)	(0.038)	(0.011)	(4.9E-07)	(3.6E-04)	(0.011)	(0.001)
R^2	0.782									
Steaks										
Coefficient	0.956*	-0.364*	-0.688*	-0.557*	-0.493*	0.345*	2.8E-05*	-6.7E-11*	-0.379*	0.020*
Robust Std. Error	(0.010)	(0.057)	(0.056)	(0.057)	(0.057)	(0.017)	(7.5E-07)	(5.6E-04)	(0.015)	(0.002)
R^2	0.746									
Other										
Coefficient	0.916*	0.031	-0.167*	-0.185*	0.176*	0.196*	9.6E-06*	-2.2E-11*	-0.126*	0.009*
Robust Std. Error	(0.017)	(0.054)	(0.053)	(0.053)	(0.054)	(0.014)	(6.5E-07)	(5.0E-04)	(0.014)	(0.002)
R^2	0.718									
Frozen										
Coefficient	1.042*	-0.118	0.458*	-0.162	0.292*	-0.133*	1.1E-05*	-3.7E-11*	-0.169*	0.006
Robust Std. Error	(0.027)	(0.116)	(0.116)	(0.115)	(0.114)	(0.051)	(1.3E-06)	(9.5E-04)	(0.025)	(0.003)
R^2	0.864									
PORK										
Ground pork										
Coefficient	1.007*	0.004	-0.286*	-0.193	-0.095	0.063*	5.8E-06*	-2.5E-11*	-0.040	0.001
Robust Std. Error	(0.050)	(0.108)	(0.109)	(0.110)	(0.107)	(0.024)	(1.1E-06)	(7.9E-12)	(0.027)	(0.004)
R^2	0.845									
Roasts and hams										
Coefficient	1.075*	-0.442*	-0.494*	-0.508*	-0.448*	0.035*	9.4E-06*	-2.7E-11*	-0.059*	0.004
Robust Std. Error	(0.019)	(0.055)	(0.054)	(0.055)	(0.055)	(0.018)	(7.7E-07)	(5.5E-12)	(0.017)	(0.002)
R^2	0.691									
Steaks and chops										
Coefficient	1.100*	-0.597*	-0.842*	-0.624*	-0.535*	0.108*	1.2E-05*	-4.0E-11*	-0.066*	-0.002
Robust Std. Error	(0.023)	(0.072)	(0.071)	(0.071)	(0.071)	(0.011)	(4.4E-07)	(3.2E-12)	(0.009)	(0.001)
R^2	0.811									
Other										
Coefficient	1.085*	-0.292*	-0.455*	-0.392*	-0.372*	0.073*	8.2E-06*	-9.8E-12	-0.046*	-0.002
Robust Std. Error	(0.017)	(0.050)	(0.048)	(0.048)	(0.049)	(0.017)	(7.4E-07)	(5.8E-12)	(0.014)	(0.002)
R^2	0.723									

Note: A * denotes coefficients that are statitically significantly different from zero at the 5 percent level.

Table 3.9 Estimated Coefficients of Censored Product Price Models

	Monthly Average Group Price	Region				Urban Area	Household Income	Household Income^2	Household Size	Household Size^2
		East	Central	South	West					
POULTRY										
Ground poultry										
Coefficient	1.100*	-0.285*	-0.302*	-0.571*	-0.102	0.096*	5.2E-06*	-8.1E-12	-0.125*	0.012*
Robust Std. Error	(0.058)	(0.133)	(0.135)	(0.134)	(0.131)	(0.032)	(9.4E-07)	(6.4E-12)	(0.024)	(0.004)
R^2	0.814									
Whole, bone-in chicken										
Coefficient	0.860*	0.155*	0.065	0.133*	0.213*	0.059*	4.0E-06*	-8.0E-12*	-0.075*	0.006*
Robust Std. Error	(0.028)	(0.038)	(0.038)	(0.038)	(0.038)	(0.007)	(2.9E-07)	(2.2E-12)	(0.006)	(0.001)
R^2	0.641									
Whole, bone-in turkey										
Coefficient	1.087*	-0.176*	-0.145*	-0.121*	-0.004	0.096*	1.4E-06*	1.2E-11*	-0.097*	0.009*
Robust Std. Error	(0.022)	(0.038)	(0.037)	(0.038)	(0.038)	(0.014)	(6.1E-07)	(4.5E-12)	(0.012)	(0.002)
R^2	0.603									
Boneless										
Coefficient	0.846*	0.346*	0.282*	0.343*	0.662*	0.049*	5.5E-06*	-8.7E-12*	-0.132*	0.008*
Robust Std. Error	(0.013)	(0.044)	(0.044)	(0.043)	(0.045)	(0.013)	(5.0E-07)	(3.4E-12)	(0.011)	(0.002)
R^2	0.811									
Other										
Coefficient	0.994*	-0.089	-0.301*	-0.286*	-0.083	0.004	1.1E-05*	-2.5E-11*	-0.091*	0.004
Robust Std. Error	(0.031)	(0.077)	(0.078)	(0.077)	(0.079)	(0.026)	(1.1E-06)	(8.4E-12)	(0.021)	(0.003)
R^2	0.627									
Frozen										
Coefficient	1.018*	-0.425*	-0.285*	-0.251*	-0.275*	-0.040*	8.6E-06*	-3.4E-11*	0.013	-0.005*
Robust Std. Error	(0.049)	(0.097)	(0.097)	(0.097)	(0.098)	(0.010)	(4.4E-07)	(3.3E-12)	(0.009)	(0.001)
R^2	0.841									

Note: A * denotes coefficients that are statitically significantly different from zero at the 15 percent level.

Table 3.10 Summary Statistics of Observed and Predicted Monthly Prices[a]

	Average	Minimum	Maximum	Std. Dev.	N
BEEF					
Ground beef					
Observed Price	2.160	0.000	36.014	0.937	279,931
All Prices[b]	2.226	0.000	36.014	0.632	745,632
Roasts					
Observed Price	2.532	0.000	36.362	1.410	127,561
All Prices	2.573	0.000	36.362	0.694	745,632
Steaks					
Observed Price	4.511	0.000	36.355	2.821	203,330
All Prices	4.596	0.000	36.355	1.654	745,632
Frozen					
Observed Price	3.142	0.000	19.980	1.373	15,366
All Prices	3.431	0.000	19.980	0.527	745,632
Other					
Observed Price	2.704	0.000	36.393	1.779	108,403
All Prices	2.787	0.000	36.393	0.798	745,632
PORK					
Roasts and hams					
Observed Price	2.042	0.000	17.447	0.932	9,855
All Prices	2.053	0.000	17.447	0.297	745,632
Steaks and chops					
Observed Price	2.427	0.000	18.136	1.784	79,925
All Prices	2.404	0.000	18.136	0.745	745,632
Frozen					
Observed Price	2.510	0.000	18.137	1.725	75,355
All Prices	2.524	0.000	18.137	0.737	745,632
Other					
Observed Price	2.932	0.000	18.030	1.474	166,784
All Prices	2.980	0.000	18.030	0.750	745,632
POULTRY					
Ground poultry					
Observed Price	2.199	0.000	20.000	1.099	23,678
All Prices[b]	2.164	0.000	20.000	0.347	745,632
Whole, bone-in chicker					
Observed Price	1.304	0.000	20.333	0.995	189,805
All Prices	1.353	0.000	20.333	0.527	745,632
Whole, bone-in turkey					
Observed Price	1.268	0.000	20.423	1.097	49,065
All Prices	1.410	0.000	20.423	0.366	745,632
Boneless					
Observed Price	2.701	0.000	19.974	1.380	125,262
All Prices	2.755	0.000	19.974	0.671	745,632
Frozen					
Observed Price	1.987	0.000	17.320	0.883	65,270
All Prices	1.981	0.000	17.320	0.304	745,632
Other					
Observed Price	2.261	0.000	20.636	1.870	40,991
All Prices	2.283	0.000	20.636	0.648	745,632

[a] Average prices differ slightly from the values in table XX due to aggregation of daily transactions to monthly level. Prices are reported in $/lb.

[b] Includes both observed and predicted prices.

Table 3.11 Summary Statistics of Quality Adjusted Monthly Purchases and Price Indices[a]

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

[a] Summary statistics based on 745,632 monthly observations.

Table 3.12 Names and Locations of Newspapers used in Media Index

Region	City	State	Newspaper Name
Central	Chicago	IL	Chicago Sun Times
	Columbus	OH	Columbus Dispatch
	Grand Rapids	MI	Grand Rapids Press
	Milwaukee	WI	Milwaukee Journal Sentinal
	St. Louis	MO	St. Louis Post-Dispatch
Northeast	Minneapolis	MN	Star Tribune
	Buffalo	NY	Buffalo News
	New York	NY	Daily News
	Pittsburgh	PA	Post-Gazette
	Boston	MA	Boston Globe
	New York	NY	New York Times
South	Newark	NJ	Star-Ledger
	Little Rock	AR	Arkansas Democrat-Gazette
	St. Petersberg	FL	St. Petersberg Times
	Atlanta	GA	Atlanta Journal
	Houston	TX	Houston Chronicle
	Washington	DC	Washington Post
West	New Orleans	LA	Times-Picayune
	Sacramento	CA	Sacramento Bee
	Denver	CO	Denver Post
	Portland	OR	Oregonian
	San Diego	CA	San Diego Union-Tribune
	San Francisco	CA	San Franciso Chronicle

Table 3.13 Summary Statistics of Monthly Household Purchase Occasions

	Average	Median	Minimum	Maximum	Std. Dev.
Beef	1.956	2	1	19	1.228
Pork	1.440	1	1	17	0.762
Poultry	1.601	1	1	18	0.931

Table 3.14 Summary Statistics of Monthly Food Safety Information, 1998 to 2005[a]

	Average	Minimum	Maximum	Std. Dev.
Newspaper Articles				
East				
Beef	14.625	9.049	1.000	51.000
Pork	4.990	3.447	0.000	22.000
Poultry	18.906	8.162	3.000	50.000
Central				
Beef	15.688	11.669	2.000	81.000
Pork	5.010	3.127	0.000	19.000
Poultry	22.865	11.004	5.000	69.000
South				
Beef	20.281	12.111	6.000	76.000
Pork	7.729	5.658	1.000	42.000
Poultry	34.604	12.317	18.000	79.000
West				
Beef	13.094	16.532	1.000	123.000
Pork	3.156	2.621	0.000	18.000
Poultry	15.010	8.374	4.000	59.000

[a] Sample size equals 96 monthly observations.

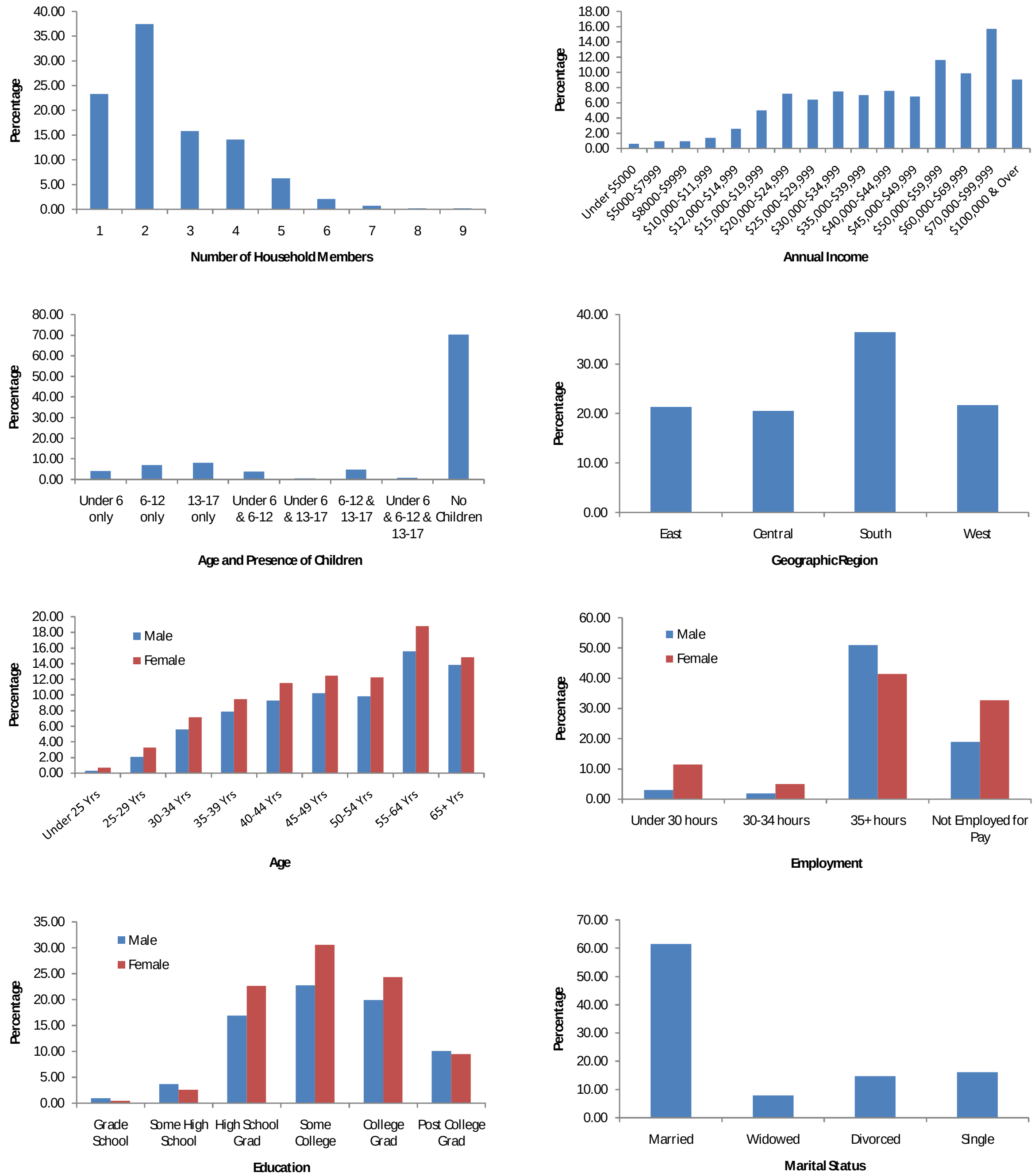

Figure 3.1 Nielsen Panel Household Characteristics

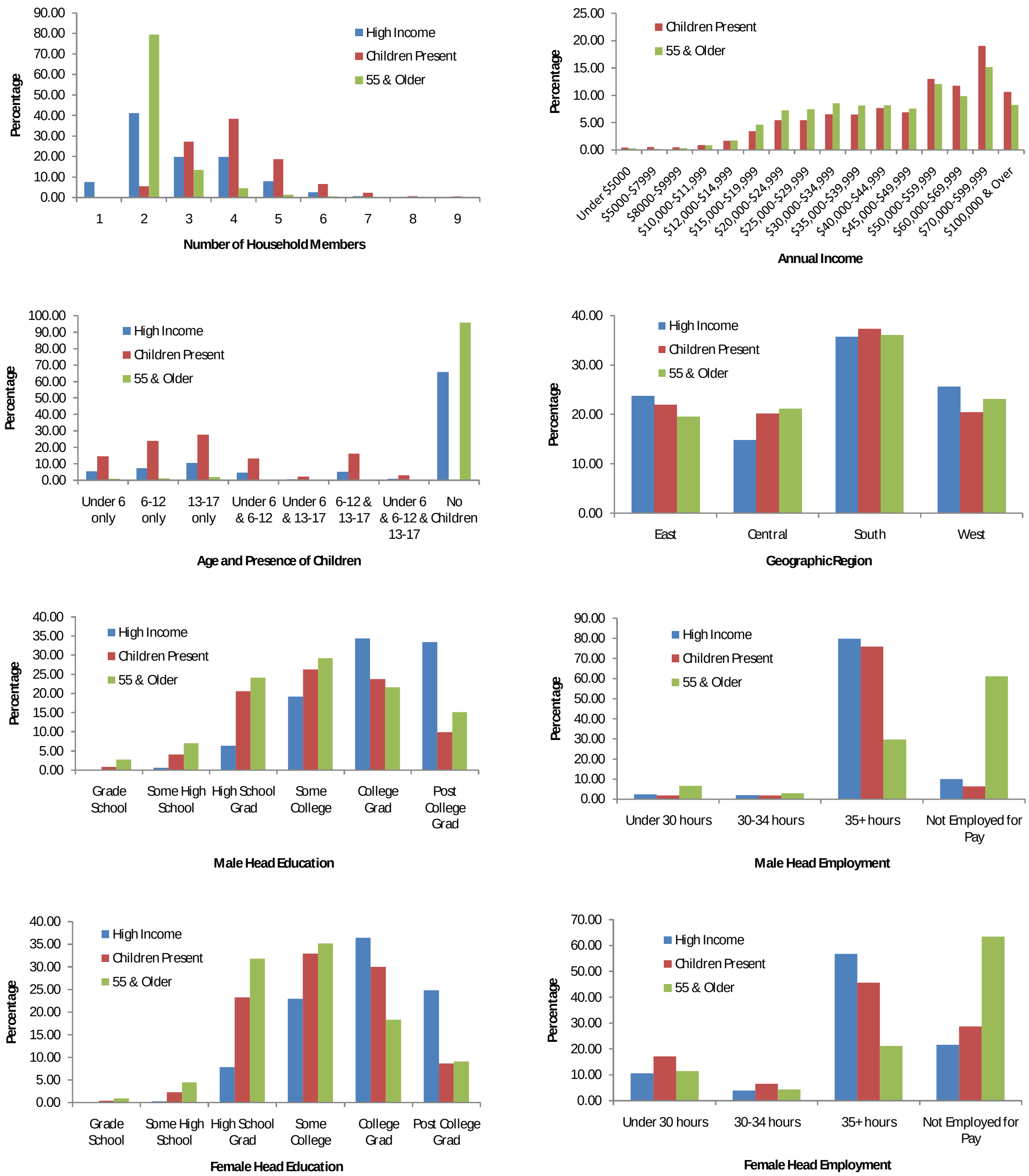

Figure 3.2 Nielsen Panel Household Characteristics by Demographic Group

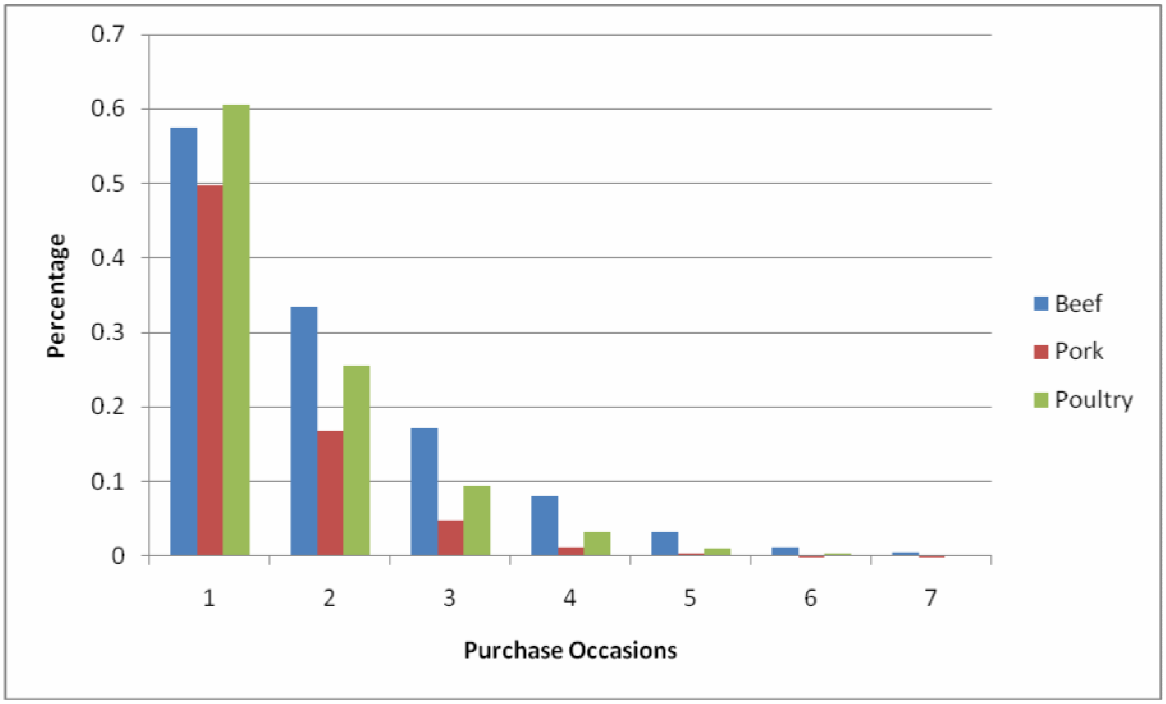

Figure 3.3 Frequencies of Monthly Household Purchase Occasions for Beef, Pork, and Poultry Products

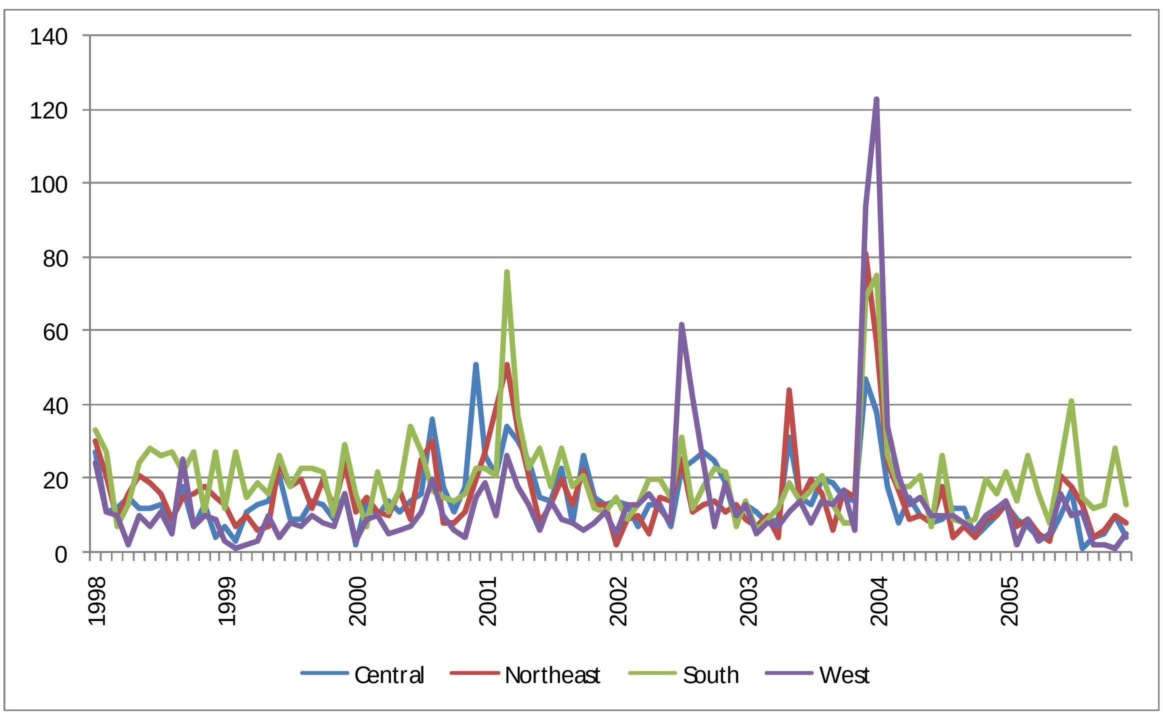

Figure 3.4 Beef Media Index by Region, 1998 to 2005

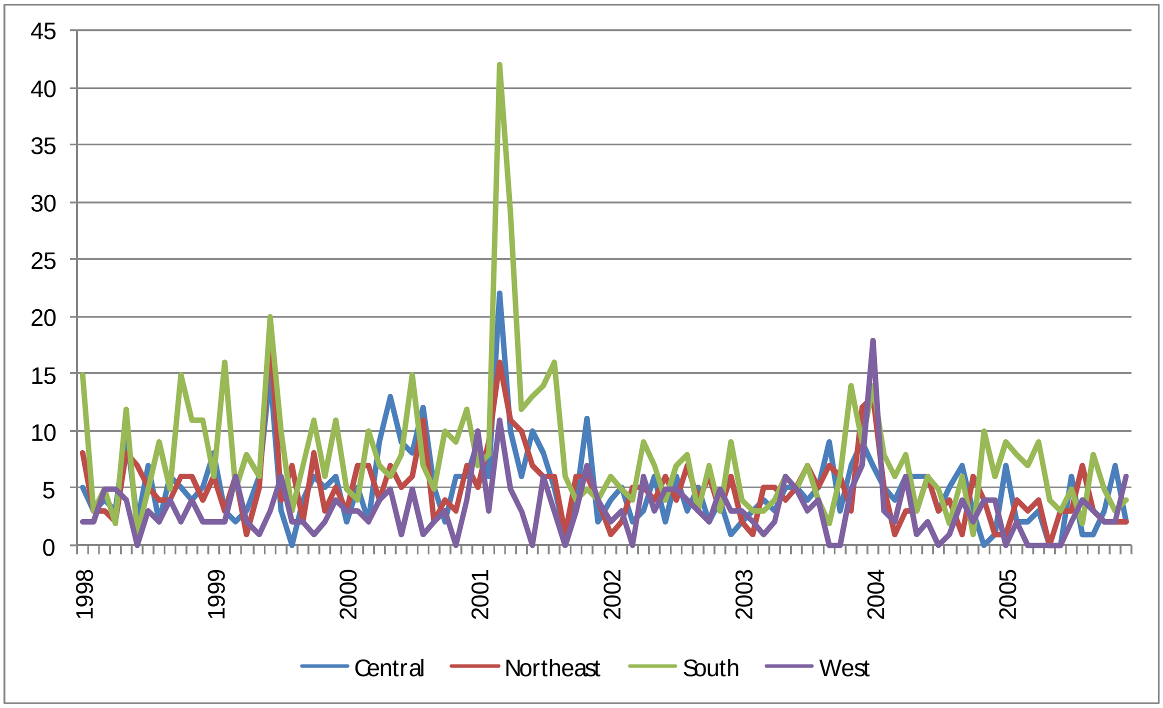

Figure 3.5 Pork Media Index by Region, 1998 to 2005

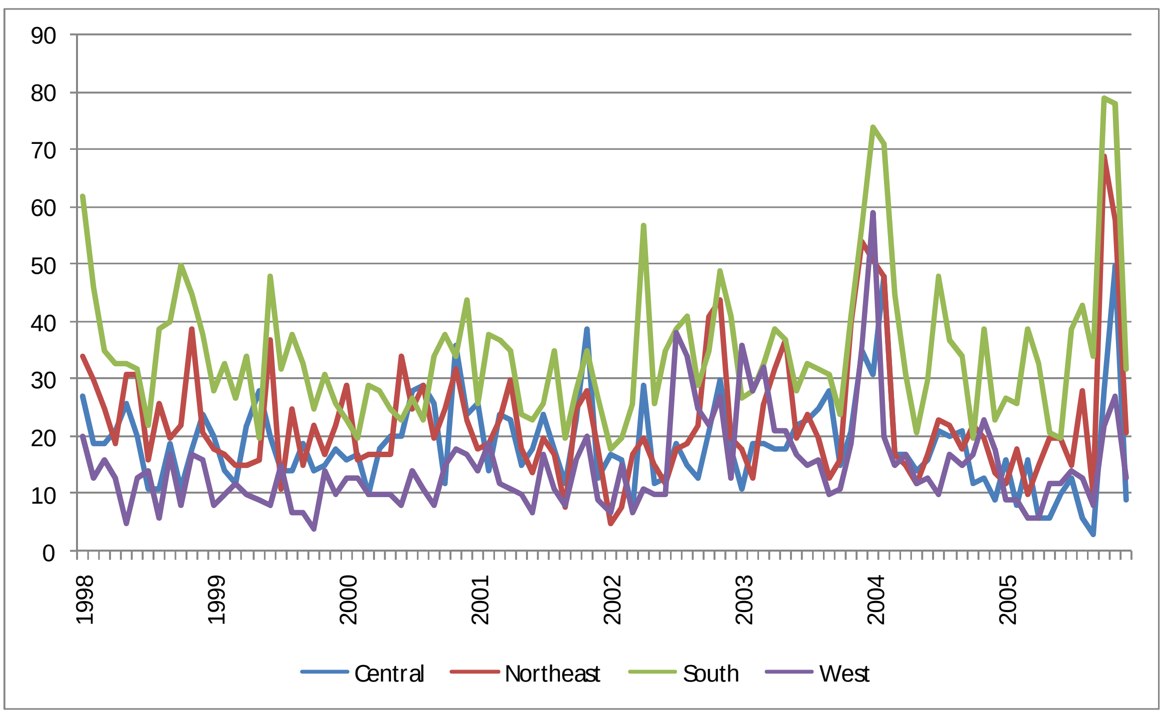

Figure 3.6 Poultry Media Index by Region, 1998 to 2005

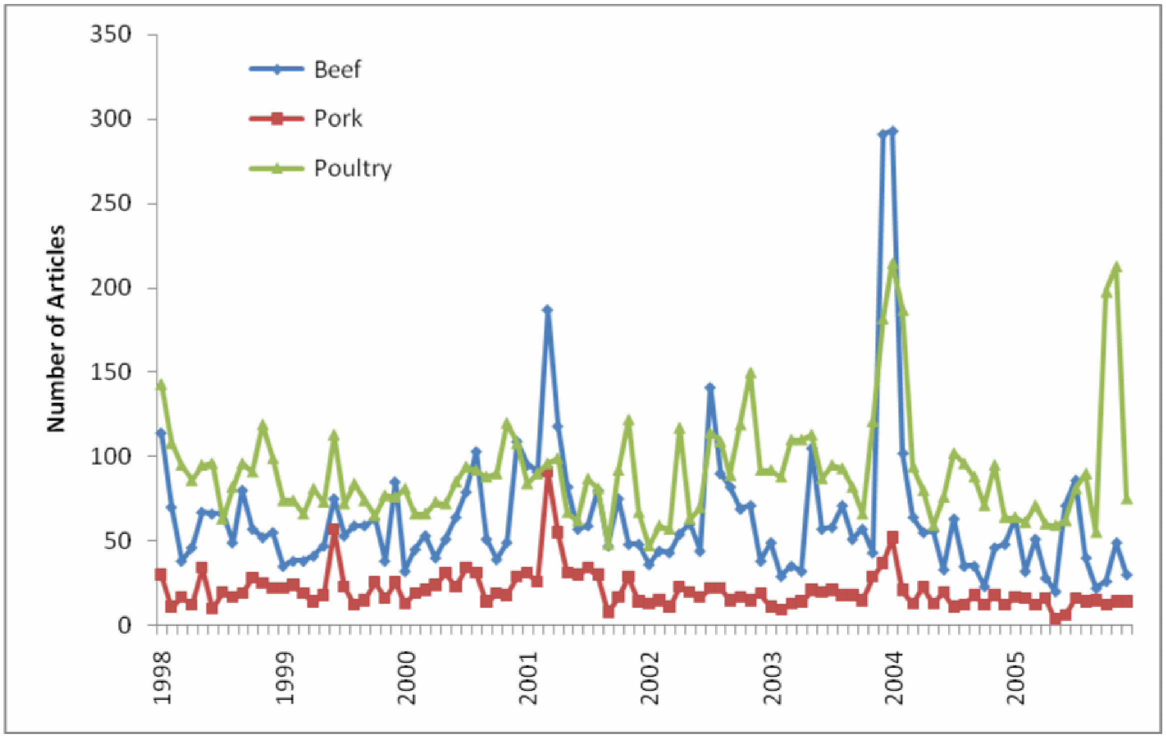

Figure 3.7 Total beef, pork, and poultry food safety articles, 1998 to 2005

4 Chapter

Discrete Choice Models of Meat and Poultry Purchases

4.1 Introduction

Many factors can influence consumer purchasing habits. Food safety concerns may include previous experience with foodborne illness, warnings of food safety risks from a physician, or receiving information on food safety in the media.[13] While idiosyncratic experiences are difficult to measure, the amount of food safety information available to consumers in the press can be quantified. 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, however it does not allow assessment of the likelihood that individual households will avoid purchasing meat and poultry products in response to food safety information. Examining this type of discrete avoidance behavior at the disaggregate level (i.e., what mix of products households buy on a

[13] 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.

given purchase occasion) will provide additional and complementary insight into the demand for food products under different food safety information environments.

The objective of the models presented in this chapter is to investigate if the quantity of food safety information publicly available impacts consumers' decisions to purchase fresh meat and poultry in a discrete choice framework. 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-specific, monthly parameters are constructed using the media index and discrete choice models of product choice are estimated to measure the impact of food safety information on purchase behavior. Results from the binary and multinomial conditional logit models will provide insight into households' propensity to avoid consumption of a commodity or substitute to another when faced with food safety concerns. A second objective of the analysis presented here is to determine which factors are likely to affect consumer behavior and use these parameters in the specification of the demand models of total quantity purchased introduced in the next chapter.

4.2 Binary Choice Models

The consumption data described in the previous chapter measures how much of each meat and poultry product each household bought in a given month. This detail allows for the purchase patterns to be modeled in a variety of ways. The first of two modeling strategies presented in this chapter is a simple binary choice model. In this model, people are recorded as either buying a good or not. For example, if a household bought beef in a given month, then the binary choice vector used in the regression has a value equal to one. If they did not buy beef, then the value is equal to zero. This type of model does not specify how much beef was bought or if any other meat or poultry products were purchased. A more complex model of purchases that further specifies any additional or alternative meat and poultry purchases by a household is presented in the next section.

4.2.1 Logit Model Derivation

The binary choice situation described above is estimated using a logit model. The derivation of the logit 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). 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:

$$\begin{aligned} P_{nj} &= \Pr\text{ob}\left(V_{nj} + \varepsilon_{nj} > V_{ni} + \varepsilon_{ni}, \forall j \neq i\right) \\ &= \Pr\text{ob}\left(\varepsilon_{ni} < \varepsilon_{nj} + V_{nj} - V_{ni}, \forall j \neq i\right) \ . \end{aligned} \tag{4.1}$$

Since ε_{nj} is not known, the probability is the integral of $P_{nj}\big|\varepsilon_{nj}$ over all values of ε_{nj} weighted by its density:

$$P_{nj} = \int \left(\prod_{j \neq i} e^{-e^{-(\varepsilon_{nj} + V_{nj} - V_{ni})}} \right) e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}} d\varepsilon_{nj} \ . \tag{4.2}$$

Some algebraic manipulation yields the following closed form expression for the logit probability of alternative *j* for individual *n*:

$$P_{nj} = \frac{e^{V_{nj}}}{\sum_{j=1}^{J} e^{V_{nj}}} \ , \tag{4.3}$$

which can be written as follows for the case of $J = 2$ alternatives as:

$$P_n = \frac{e^{V_n}}{1 + e^{V_n}} \ . \tag{4.4}$$

The binary choice model is estimated individually for each of three alternatives (beef, pork, and poultry). In each model, V_n is specified as a linear function of the parameters. Therefore, equation (4.4) can be written as:

$$P_n = \frac{e^{\boldsymbol{\beta}'\mathbf{x}_n}}{1 + e^{\boldsymbol{\beta}'\mathbf{x}_n}} \ , \tag{4.5}$$

where $\mathbf{x}_n$ is a vector of alternative-specific and person-specific characteristics and $\boldsymbol{\beta}$ is a corresponding vector of estimated coefficients. The log-likelihood function used in model estimation is as follows:

$$\ln \underset{\max \beta}{L}(\boldsymbol{\beta}) = \sum_{i=1}^{N} \left\{ \mathbf{y}_i \ln(P_n) + (1 - \mathbf{y}_i) \ln(1 - P_n) \right\} , \tag{4.6}$$

where $\mathbf{y}_i$ is an indicator vector with value equal to one if the single alternative is chosen and zero otherwise.

4.2.2 Model Specification

Three binary choice models are estimated for each of the three commodities of interest; beef, pork, and poultry. The models are comprised of characteristics of the alternatives as well as those of the household. The models also contain monthly binary variables and interaction terms between the food safety information and select demographic variables. The binary choice model for commodity *k* is a linear function of parameters with the following specification:

$$V_n^k = \sum_{k=1}^{3} \gamma_k Price_{nk} + \sum_{k=1}^{3} \beta_k MI_{nk} + \eta Ed_n * MI_{nk} + \delta Age_n * MI_{nk} \tag{4.7}$$
$$+ \mu Urban_n * MI_{nk} + \rho Child_n * MI_{nk} + \sum_{d=1}^{D} \tau^d h_n^d \ ,$$

where *D* indexes the total number of demographic variables included in the model, and h_n^d is the *d*th demographic characteristic of household *n*, and *k* indexes the three commodities. Summary statistics of the variables used in each of the three binary choice models are listed in table 4.1.

The variable *Price* used in the three binary choice models 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 each month. 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.[14] *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 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

[14] 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.

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.[15] 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 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

[15] 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.

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.

4.2.3 Estimation and Results

The binary choice logit model was estimated using the statistical software STATA. The full dataset contains 745,632 monthly household observations, however the use of two week lagged food safety information to construct the media index variables in each of the models required 7,465 observations from the first month of 1998 to be dropped from the subsample. The summary statistics of the full set of observations are listed in table 4.1.

Within the final subsample of monthly household observations are 17,428 unique households. It is reasonable to assume that some correlation between observations from the same household may exist. Therefore, clustered robust standard errors are estimated using the unique households in the panel. The clustered robust standard error is based on the robust sandwich estimator (Huber, 1967; White, 1980) and is specified as:

$$V_{robust} = \left(X'X\right)^{-1} * \left[\sum_{i=1}^{n}\left(e_i * \mathbf{x}_i\right)' * \left(e_i * \mathbf{x}_i\right)\right] * \left(X'X\right)^{-1} \quad , \tag{4.8}$$

where X is the matrix of regressors, e_i is the residual for the ith observation, $\mathbf{x}_i$ is a row vector of regressors, and n is the total number of observations. The clustered robust standard error uses the following adjustment to the robust estimator:

$$V_{cluster} = \left(X'X\right)^{-1} * \sum_{c=1}^{n_c} u_c' * u_c * \left(X'X\right)^{-1} \quad , \tag{4.9}$$

where $u_c = \sum_{i \in c} e_i * \mathbf{x}_i$ and n_c is the total number of clusters. In this model, n_c equals the number of unique households and allows for correlation of observations within, but not between households. The non-normality of the errors results in the model being estimated using maximum pseudolikelihood. This estimation technique maximizes the same log-likelihood function described in equation (4.6), but requires the estimated asymptotic covariance to be adjusted as described above.

4.2.3.1 Beef Purchase Models

The results of the binary choice model for the purchase of beef are presented in table 4.2. The parameter for the own-effect of the price of beef has the expected negative sign and is statistically significantly different from zero. The coefficients for the effects of the price of pork and poultry on beef purchase are also negative and statistically significant. The signs of the cross-price effects indicate that pork and poultry are complements to beef, which does not correspond with the a priori expectation that they are substitutes.

The coefficient for the regional media index variable, *MI*, is not statistically significantly different from zero. This coefficient indicates that there is no effect from an increase in food safety information related to beef on the probability of households making monthly purchases of fresh beef. The beef media index interaction terms with select demographic variables indicate that certain groups of household do, however, respond to changes in food safety information. Households with children have a decreased probability of purchasing beef when the beef media index increases. This effect is statistically significant at the 5 percent level. The interaction terms for college education and heads of household aged 55 and older were also negative in sign, but are not statistically significantly different from zero. The interaction term for households located in urban areas is also not statistically significantly different from zero.

The parameters for *Income* and *Income*^{2} are both statistically significantly different from zero. The positive sign of *Income* indicates that increases in total household income will increase the probability of purchasing beef on a monthly basis. The negative sign for *Income*^{2} indicates that the positive effect from income declines as income increases. The coefficient for *Hsize* is statistically significantly different from zero and has a negative sign indicating that the probability of purchasing beef declines for larger households.

Seasonal effects on the probability of purchasing beef are included in the model as monthly binary variables. All but one of the monthly parameters included in the model are statistically significant at the five percent level. The parameters are positive for all months except November, indicating that the people are less likely to purchase beef in November and December relative to the rest of the year. Most of the year dummy variables are statistically significant and indicate a year to year fluctuation in purchase probability that does not necessarily follow a trending pattern over the time period analyzed.

The coefficients for the geographical location of a household are interpreted relative to the region that is omitted from the regression. In this model that is the southern region of the United States. The coefficients for *Central* and *Northeast* are not statistically significantly different from zero, indicating that households in the central and northeastern regions are no more or less likely to purchase beef than households in the southern region. Based on the positive and statistically significant *West* parameter, households in the western region of the country are more likely to purchase beef on a monthly basis than those in the south.

The binary variables denoting the race of the head of household are interpreted relative to the omitted race variable, *Caucasian*. None of the race variables are statistically significantly different from zero. This indicates that none of these groups are more or less likely to purchase beef than Caucasian households.

4.2.3.2 Pork Purchase Models

The estimates of the binary choice model of monthly pork purchases are listed in table 4.2. The coefficient for pork price has the expected negative sign and is statistically significantly different from zero. The coefficients for the cross-effects of the price beef and poultry on the probability of making a pork purchase are negative and statistically significant. As with the beef model, this is an unusual sign given that beef and poultry are considered substitute goods for pork.

The estimated coefficient for the regional media index of pork food safety articles is statistically significantly different from zero, but has an unexpected positive effect on the probability of purchasing pork. Although the average effect of pork food safety media has an unexpected positive sign, the signs of the interaction terms for food safety information are negative. The signs of interaction terms for households with heads aged 55 and older as well as households with children indicate that these groups respond negatively to additional pork food safety media. The coefficients for these interactions are statistically significant at the 1 and 2 percent levels, respectively. The coefficients for households with a college educated head and for households located in urban areas are also negative, but not statistically significant. These results indicate that differences in the effect of food safety information on pork purchase decisions exist for households with varying characteristics.

The estimated coefficients for *Income* and $Income^2$ are both statistically significantly different from zero. The positive sign of *Income* indicates that increases in total household income will increase the probability of purchasing pork, while the negative sign for $Income^2$ indicates that the positive income effect declines as income increases. The coefficient for household size is not statistically significantly different from zero, indicating that changes in family size do not impact the probability of purchasing pork on a monthly basis.

Each of the monthly parameters included in the model are statistically significant at the five percent level and negative in sign. The parameters suggest that pork purchases are higher in December relative to the rest of the year. All but one of the year dummy variables are statistically significant and indicate a year to year fluctuation in purchase probability that does not follow a trend pattern.

The coefficients for the *Northeast* and *Central* regions are not statistically significantly different from zero, indicating that households in the northeastern and central regions are no more or less likely to purchase pork than households in the southern region. The *West* parameter is statistically significantly different from zero at a 5 percent level and the positive sign suggests that households in the western region of the country are more likely to purchase pork on a monthly basis than those in the south.

The binary variable denoting an Asian head of household is statistically significantly different from zero at a 5 percent level. The negative sign indicates that these households are less likely to purchase pork on a monthly basis than their Caucasian counterparts. All other race variables are not statically significant, indicating that these groups are no more or less likely to make monthly pork purchases than Caucasian households.

4.2.3.3 Poultry Purchase Models

The estimates of the binary choice model of monthly poultry purchases are listed in table 4.2. The coefficient for poultry price has the expected negative sign and is statistically significantly different from zero. The coefficients for the cross-effects of the price beef and pork on the probability of purchasing poultry are also negative and statistically significant. As with the beef and pork models, this is an unusual sign given that beef and pork are considered substitute goods for pork.

The estimated coefficient for the regional media index of poultry food safety articles is not statistically significantly different from zero. This suggests that changes in the amount of food safety information related to poultry do not, on average, affect households' decisions to make monthly purchases of poultry. Likewise, the interaction terms between the poultry media index and select demographic variables suggest non-response of consumers to poultry food safety media because none of them are statistically significantly different from zero at the 5 percent level.

The estimated coefficient for *Income* is statistically significantly different from zero at the 1 percent level, as it is in both the beef and pork models. The positive sign of *Income* indicates that increases in total household income will increase the probability of purchasing poultry. The $Income^2$ parameter is also statistically significantly different from zero and the negative sign indicates the positive effect of income diminishes as income increases. The coefficient for household size is statistically significant and negative in sign, indicating that an increase in the size of the household decreases the probability of purchasing poultry on a monthly basis.

Each of the monthly parameters included in the model are statistically significant at the five percent level and positive in sign. The parameters suggest that poultry purchases are lower in December relative to the rest of the year. All of the year dummy variables are statistically significant and, as with beef and pork, indicate a year to year fluctuation in purchase probability that does not follow a trend pattern over the sample period.

The coefficients for the *West* and *Northeast* regions are statistically significantly different from zero at the 1 percent level. The positive signs of the coefficients indicate that households in the western and northeastern regions are more likely to purchase poultry than households in the southern region. The *Central* parameter is not statistically significantly different from zero. This suggests that households in the central region of the country are no more or less likely to purchase poultry on a monthly basis than those in the south.

The binary variables denoting the race of the head of household are not statistically significantly different from zero. These parameters indicate that none of the groups are more or less likely to purchase beef than Caucasian households.

4.2.4 Binary Choice Model Elasticities

Elasticities were calculated for price and income effects across the entire population as well as food safety effects for the various demographic subgroups. The statistical significance of the individual parameters does not reveal the total effect of food safety information on the probability of purchasing meat and poultry for the demographic groups considered in the model. The elasticities, however, are calculated using both the average and interacted effects, thereby measuring a total effect. Elasticities are also unitless measures, which allows for a comparison of price and income effects relative to food safety information.

In addition to calculating a mean elasticity effect for the parameters of interest (price, income, and food safety), the Krinsky-Robb simulation technique is employed to generate empirical distributions of the parameters (Krinsky and Robb, 1991).[16] The simulation technique involves drawing realizations of the model parameters from a multivariate normal distribution with a mean and covariance matrix of the estimated model parameters as follows:

$$\beta^{KR} \sim N\left(\boldsymbol{\beta}, S\right) , \tag{4.10}$$

where β^{KR} is a $\left(p \times r\right)$ matrix of parameter realizations, $\boldsymbol{\beta}$ is a $\left(1 \times r\right)$ vector of estimated parameter coefficients, S is the $\left(r \times r\right)$ estimated variance-covariance matrix, and p is the

[16] The Krinsky-Robb simulation technique is a non-parametric bootstrap method. A parametric bootstrap, which provides the parameter distribution, may be preferred if a normal distribution is a not an accurate representation of the parameter distribution.

number of realizations drawn for the simulation.[17] Following Krinsky ad Robb (1986), 1000 draws are taken for this simulation. With the parameter draws in hand, elasticities of the select parameters are calculated for every simulated parameter realization. The resulting empirical distribution of elasticities can be summarized to determine statistical properties of the elasticities such as standard deviations and confidence intervals.

The logit elasticity is the percentage change in the probability of choosing an alternative for a one percent change in another variable. It is calculated for commodity *k* with respect to parameter x_n as follows:

$$E_{x_n}^k = \frac{\partial P_n^k}{\partial x_n}\frac{x_n}{P_n^k} = \frac{\partial V_n^k}{\partial x_n} x_n \left(1 - P_n^k\right) , \tag{4.11}$$

where P_n^k is the probability of household *n* choosing commodity *k*. The elasticities reported in table 4.3 are the means of all the individual household elasticities. For the price and income elasticities, this average is taken over all the households in the sample. The food safety elasticities, however, are calculated using only the households included in each of the four demographic groups considered.

In each of the beef, pork, and poultry models, the own-price elasticities are statistically significantly different from zero using a 95% confidence interval. The elasticity for beef with respect to *Price* indicates that a 1% increase in the price of beef decreases the probability of purchasing beef by 3.42%. Similarly, the effect of a 1% increase in the price of poultry is estimated to be a decrease of 3.99% of the probability of purchase. The own-price

[17] The Krinsky-Robb technique appeals to the asymptotic normality of the parameter vector. The parameters of the logit model are estimated using maximum likelihood. Asymptotic normality of the parameter vector is one of the properties of the maximum likelihood estimator. Therefore, the Krinsky-Robb simulation method is appropriate for the parameters of the logit model.

elasticity of pork is smaller in magnitude as compared to beef and poultry, with the probability of purchase declining by 1.87% for a 1% increase in the price of pork. The cross-price elasticities for each of the three commodities are also statistically significant at a 95% level and negative in sign. However, these effects are much smaller than the own-price effects with magnitudes ranging from 0.06% to 0.57%. The elasticities for the probability of purchasing beef, pork, and poultry with respect to income are each statistically significant and positive in sign. A 1% increase in income increases the probability of purchasing beef, pork, and poultry by .47%, 0.34%, and 0.51%, respectively.

All but three of the food safety elasticities for the four demographic groups are not statistically significantly different from zero at a 95% level. This suggests that, for most of the households considered in the model, food safety information does not have a measurable impact on the probability of purchasing beef, pork, and poultry on a monthly basis. For the households that do have a statistically significant response, the effect is small in magnitude relative to the effect of prices or income. The beef food safety elasticity for households with children is statistically significantly different from zero and suggests a 0.03% decrease in the probability of purchasing beef from a 1% increase in the beef food safety index. The other food safety elasticities that are statistically significant are the college education and urban location elasticities from the pork model. These elasticities have an unexpected positive sign, with an estimated 0.03% increase in the probability of purchasing pork on a monthly basis from a 1% increase in the pork food safety information index.

4.2.5 Binary Choice Model Summary

Binary logit models are simple models of consumer behavior. Households may choose to buy a given commodity or not, but other alternatives are not available. However, simple models can provide insight into the basic interactions of purchase decisions with variables such as price, income, and food safety information. Results from these models indicate that food safety information, as measured by a regional media index, does impact household purchases of meat and poultry. The elasticities of the logit probabilities provide some evidence of household-level heterogeneity being a determinant in consumer response to food safety information. The effect on beef purchases from increased food safety information is negative for households with children, but the effect on the probability of purchasing pork is positive for college educated and urban households in response to higher levels of pork food safety information. The unexpected signs of the pork elasticities could be due to model misspecification. Therefore, further empirical investigation is conducted to determine if a more complex choice set may be a more accurate specification of household purchase decisions for meat and poultry.

4.3 Multiple Choice Models of Meat and Poultry Purchases

The binary choice models presented in the previous section, while a simplified representation of the consumer's choice set, are useful models for revealing some of the factors that affect the probability of making a purchase. However, the data available for this study allow for further investigation of purchase patterns. It is not only known if a household bought beef in a given month, but also if that household bought pork, poultry, all three meats, or none of them. Incorporating this information into a multinomial choice model will allow for any interactions among the three commodities and reveal the probabilities of a household purchasing each of the goods as well as combinations of them.

4.3.1 Alternative-Specific Logit Model Derivation

A logit model is again specified for this estimation procedure. The model is similar in its derivation to the binary choice model described in section *4.2.1*. Adjustments are made for multiple alternatives as well as alternative-specific constants.

The conditional logit model with alternative-specific constants, is motivated by a random utility model where the *n*th household faces *J* alternatives and the utility of alternative *j* is:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \ . \tag{4.12}$$

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

$$V_{nj} = \alpha_j + \boldsymbol{\beta}_n'\mathbf{x}_n + \boldsymbol{\beta}_j'\mathbf{x}_j \tag{4.13}$$

where α_j is an alternative-specific constant term for alternative *j*, $\mathbf{x}_n$ is a vector of characteristics describing household *n*, $\mathbf{x}_j$ is a vector of characteristics specific to alternative *j*, and the corresponding vectors of estimated coefficients are $\boldsymbol{\beta}_n$ and $\boldsymbol{\beta}_j$. 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:

$$P_{nj} = \frac{e^{V_{nj}}}{\sum_{j=1}^{J} e^{V_{nj}}} \quad . \tag{4.14}$$

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:

$$\ln_{\max \beta} L(\boldsymbol{\beta}) = \sum_{n=1}^{N} \sum_{j=1}^{J} \mathbf{d}_{nj} \ln P_{nj} \ , \tag{4.15}$$

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

The multinomial logit model implies a proportional substitution pattern across alternatives (Train, 2003). This property arises given due to the specification of an independent error distribution (iid). This property of the logit model is referred to as *independence from irrelevant alternatives* or IIA. The use of a model with the IIA property is not restrictive so long as the relative probability of choosing one alternative over another is the same, regardless of the other alternatives available or the attributes of the other alternatives. However, IIA may not always be an accurate representation of the true substitution patterns underlying the data. In these cases, the degree to which the IIA property restricts choice behavior will determine the amount of error generated by policy counterfactuals based on the logit model.

In the current work, IIA is only evident in the cross-elasticities of the logit probabilities. The elasticity formula for the probability of alternative *j* with respect to a variable that enters the representative utility of alternative *i* is:

$$E_{jx_{ni}} = -\beta_x x_{ni} p_{ni} \ , \tag{4.16}$$

where p_{ni} is the probability of individual *n* choosing alternative *i*. Since *j* does not enter the formula, a change in alternative *i* will affect the probabilities for all the other alternatives by the same proportion. As a result of IIA, the cross-elasticities from the multinomial logit model are unlikely to be an accurate representation of the true cross-elasticities and are not reported in this study.

4.3.2 Model Specification

The multinomial conditional logit model is estimated using a choice set of eight different alternatives. The eight purchase alternatives a household faces in a given month 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 (4.10), 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:

$$V_{nj} = \alpha_j + \sum_{k=1}^{3} \gamma_k Price_{nk} I_k^j + \sum_{k=1}^{3} \sum_{l=1}^{2} \nu_{lk} Price_{nl} I_l^j + \sum_{k=1}^{3} \beta_k M_{nk} + \sum_{k=1}^{3} \sum_{l=1}^{2} \phi_{lk} M_{nl} + \sum_{k=1}^{3} \eta_k Ed_n * M_{nk} + \sum_{k=1}^{3} \delta_k Age_n * M_{nk} + \sum_{k=1}^{3} \mu_k Urban_n * M_{nk} + \sum_{k=1}^{3} \rho_k Child_n * M_{nk} + \sum_{d=1}^{D} \sum_{k=1}^{3} \tau_k^d h_{nk}^d I_k^j \ , \tag{4.17}$$

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, I_l^j is an indicator function that is equal to 1 if commodity $l \in$ the j^{th} alternative and equal to 0 otherwise, h_{nk}^d is the d^{th} demographic characteristic of household *n* for commodity *k*, *d* indexes the total number of demographic variables included in the model, *k* and *l* each index the three commodities of interest, and *j* indexes the eight alternatives. The own-effect media index parameter, M_{nk}, is the interaction of the commodity- and region-specific media index variable for household *n* and the indicator function ($MI_{nk} * I_k^j$). This variable is the value of the media index for commodity *k* if the indicator function equals 1 for commodity *k* and equal to 0 otherwise. The cross-effect media index parameter, M_{nl}, is similarly defined as the interaction of the media index variable for household *n* and the indicator function ($MI_{nl} * I_l^j$). It equals the value of the media index for commodity *l* if the indicator function equals 1 for commodity *k* and 0 otherwise. All remaining parameters are defined in section *4.2.2*.

With the exception of the alternative-specific constants, the parameters 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. This bundling specification alters the interpretation of the coefficients relative to the binary choice model. The estimated coefficient for the commodity-specific price coefficient, γ_k, can be interpreted as the effect of the price of the k^{th} commodity 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. The estimated coefficients for the media index, β_k, are interpreted as the effect of additional food safety articles pertaining to commodity *k* on the probability of purchasing that commodity. The interpretation of the cross-media index coefficient, ϕ_k, is the effect of an increase in the media index of commodity *l* on the probability of making a purchase that includes commodity *k*.

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.

The model includes characteristics of the households that do not vary over the alternatives in the choice set, such as income, race, geographic location of the household, and seasonal effects (which are specific to the time period rather than the household, but still do not vary over alternatives). The model can be estimated such that a coefficient for each of these variables is estimated for each alternative. This would result in estimates of seven different coefficients for the effect of household income on the probably of purchasing each of those alternatives (1 through 7) relative to not purchasing beef, pork, or poultry (alternative 8). Two reasons to consider an alternative to this modeling strategy arise. First, the number of estimated coefficients increases eight-fold with each additional household characteristic considered in the model. If degrees of freedom are a consideration, it may be important to reduce the number of variables estimated. Second, and more relevant for this study, the insight from the coefficients of the individual alternatives may not be as intuitively appealing as grouping the effects into beef, pork, or poultry subsets. Therefore, commodity-specific coefficients are estimated for each household characteristic in the model. The k^{th} commodity-specific coefficient, τ_k^d, is interpreted as the effect of the d^{th} household characteristic on the probability of making a purchase that includes commodity *k*.

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 parameters 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.

4.3.3 Estimation and Results

The 8-choice logit models were estimated using the statistical software STATA. Computer limitations were met when attempting to estimate these models using the full dataset of 745,632 monthly household observations. It was determined that a sample of 3,000 households from the panel is the largest sample that can be used for estimation of the 8-choice model. Therefore, to determine the sensitivity of the model results to the particular sample used, three random samples were drawn from the full dataset for estimation purposes. The construction of the media index variables using two week lagged food safety information requires observations from the first month of 1998 to be dropped from each subsample. The summary statistics of both the full sample and the three random subsamples are listed in table 4.4. Model estimation was conducted using each of the three samples and results are presented in tables 4.5 for comparison. However, discussion of the results in the following section is limited to the estimates using the first random sample.

There are 3,000 unique households represented in the random sample drawn for estimation. As with the binary choice models, it is reasonable to assume that some correlation between observations from the same household may exist. Therefore, clustered robust standard errors are estimated using the number of unique households in the panel and estimation is done using maximum pseudolikelihood.

The results of the eight-choice regional media index model, using random sample 1 are listed in the third and fourth columns of table 4.5. The price coefficients for beef, pork, and poultry all have the expected negative sign and are statistically significantly different from zero using a 95% confidence interval. 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 significantly different from zero. 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 positive signs indicate that the beef and poultry are substitutes for pork, which is a more intuitive result than the negative cross-price coefficients estimated using the binary choice models.

In general, the multinomial logit model results indicate that changes in food safety information, as measured by the regional media indices, do not have a statistically significant impact on the probability of purchasing beef, pork, and poultry on a monthly basis. The one exception to this is the coefficient of the interaction of the beef media index and college educated heads of household. These households have a negative response to increases in the beef media index.

Several of the household demographic parameters included in the model are statistically significantly different from zero. The households that are less likely to buy fresh beef and pork on a monthly basis are those with college educated heads and those with children present. However, households with heads age 55 and older are more likely to buy fresh beef and pork. Households in urban areas are more likely to purchase fresh poultry, relative to households in rural areas. The estimated coefficient for the effect of total household income is statistically significantly different from zero and has a positive sign for beef, pork, and poultry. The quadratic income parameter has a negative sign and is statistically significant for beef and pork. The opposite signs of the income parameters indicate that an increase in total household income will increase the probability that a household will purchase meat and poultry in a given month, but that effect tapers off for beef and pork as total household income increases. The effect household size has on the probability of purchase is positive and statistically significant for beef and pork, but not poultry.

The annual and monthly parameters were included in the model to control for year- and month-specific effects not otherwise specified in the model. The vast majority of these parameters are statistically significantly different from zero at the five percent level, indicating that time and season effects are important determinants in the probability of purchasing meat and poultry.

The parameters for regional effects (*Central*, *West*, *Northeast*) vary in sign and statistical significance across the three commodities. None of the regional parameters for beef were statistically significantly different from zero, which indicates that households in the west, central, and northeast regional are no more or less likely to purchase beef than households in the southern region. Household located in the western region are less likely to purchase pork, relative to households in the southern region. There is no statistically significant difference between households in the central and northeastern region and those located in the southern region. All of the regional coefficients for poultry were statistically significantly different from zero. Households located in the central region are less likely to purchase poultry than households in the southern region, while households in the western and northeastern regions are more likely to purchase poultry.

The estimated parameters for a Hispanic head of household indicate that these households are not statistically different from Caucasian households with regard to the probability of purchasing beef or pork. They are statistically significantly more likely to purchase poultry than Caucasian households. The coefficients for black heads of household are statistically significantly different from zero for beef, pork, and poultry. The signs of the coefficients indicate that these households are less likely to buy beef and more likely to buy pork or poultry than Caucasian households. The estimated parameters for Asian heads of household are statistically significantly different from Caucasian households for beef and pork, but the coefficient for poultry is not statistically significant. Asian heads of household are less likely to purchase beef and more likely to purchase pork than Caucasian households. Both the beef and pork parameters for the *Other* race are statistically significantly different from zero. The signs of the coefficients indicate that these households are less likely to purchase beef and more likely to purchase pork than Caucasian households.

The estimated coefficients for the alternative-specific constants are all statistically significantly different from zero at the 1 percent level and have a positive sign, except the parameter for the second alternative of purchasing pork only. The positive signs of these coefficients indicate that the average effect from non-included factors on the probability of households purchasing any of these combinations of meat and poultry is positive relative to purchasing none at all.

4.3.4 Multinomial Logit Elasticities

Elasticities were calculated for price and income effects across the full sample as well as food safety effects for the various demographic subgroups. As with the binary choice models, the elasticities are calculated using both the average and interacted effects. This provides a unitless measure of the total food safety effect. The Krinsky-Robb simulation technique is again employed to generate empirical distributions of the parameters and, subsequently, the elasticities of interest.

The use of a model specification that employs bundling of alternatives results in the estimation of commodity-specific parameters. Therefore, the price, income, and food safety elasticities can be calculated using a similar commodity-specific approach. The logit elasticities are calculated for commodity *k* with respect to the parameter x_{nk} as follows:

$$E_{kx_{nk}} = \frac{\partial P_{nk}}{\partial x_{nk}} \frac{x_{nk}}{P_{nk}} = \frac{\partial V_{nj}}{\partial x_{nk}} x_{nk} \left(1 - P_{nk}\right) \ , \tag{4.18}$$

where $P_{nk} = \sum_{j=1}^{J_k} p_{nj}$, J_k is the number of alternatives that include commodity *k*, and p_{nj} is the probability of household *n* choosing alternative *j*. The elasticities reported in table 4.6 are the means over all the households in the sample for the price and income elasticities. The food safety elasticities are calculated using only the households included in each of the four demographic groups considered in the model.

The own-price elasticities for beef, pork, and poultry are each statistically significantly different from zero using a 95% confidence interval. The elasticity for beef with respect to *Price* indicates that a one percent increase in the price of beef decreases the probability of purchasing beef by 1.31%. Similarly, the effect of a 1% increase in the price of pork is estimated to decrease the probability of purchase by 1.38%. The own-price elasticity of poultry is larger in magnitude that the own-price elasticities for beef and pork, with an estimated 2.44% decrease in the probability of purchase for a 1% increase in the price of poultry. The elasticities for the probability of purchasing beef, pork, and poultry with respect to *Income* are each statistically significant and positive in sign. An income increase of 1% is estimated to increase the probability of purchasing beef, pork, and poultry by 0.14%, 0.15%, and 0.29%, respectively.

Two of the food safety elasticities for the four demographic groups are statistically significantly different from zero at the 5 percent level. For households with college educated heads, there is an estimated 0.03% decrease in the probability of purchasing beef from a 1% increase in the beef food safety index. The effect of a 1% increase in the poultry food safety index is estimated to be 0.07% decline in the probability of purchasing poultry for households located in urban areas. These effects, while statistically significant, are small in magnitude as compared to the price and income effects. The remaining household-specific food safety elasticities are not statistically significantly different from zero.

4.3.5 8-Choice Model Summary

A multinomial conditional logit model was estimated to expand on the insight gained from the binary choice models. By incorporating information from a full choice set of different meat and poultry purchase combinations, the model allows for any interactions between the three commodities and reveals the probabilities of a household purchasing each of the goods as well as combinations of them.

The estimated coefficients and resulting elasticities of the 8-choice media index model indicate that general food safety information does affect the probability of monthly household purchases of meat and poultry. Specifically, the elasticities indicate that households with college educated heads have a negative response to beef food safety information, while households located in urban areas respond negatively to increases in poultry food safety information. These results provide some evidence that there is a heterogeneous effect on the probability of purchasing beef and poultry on a monthly basis from food safety information across the households considered in this study.

4.4 Conclusion

The objective of the models presented in this chapter was to investigate if the quantity of food safety information available to consumers impacts their purchase decisions for fresh meat and poultry in a discrete choice framework. The measure of food safety information used in the models is a commodity- and region-specific media index, which represents the general presence of food safety information available to the public in their regional newspapers. The media index was modeled as interactions with various demographic characteristics to determine if the effect of food safety information varies across different groups of households.

Binary logit models were estimated to investigate the effects of the different types of food safety information on purchase decisions. While these models are rather restrictive in the specification of the choice set, they do account for the effects of price, income, food safety information, and interactions between food safety information and household demographic variables on the probability of purchasing a given commodity. Results from estimation of the commodity-specific models suggest that responses to food safety information do vary across households for beef and pork, while poultry purchase probabilities are not affected.

There were a few unexpected results from the binary choice models. First, the signs of the cross-price effects were negative, suggesting meat and poultry commodities are compliments rather than substitutes as was expected a priori. Other unexpected results included the positive signs of the pork food safety parameter and elasticities for college educated and urban households. These results prompted further investigation of the meat and poultry purchase decisions in a more complex model accounting for the interactions between purchase alternatives.

To further investigate the interactions between beef, pork, and poultry purchase decisions in the presence of food safety information, multinomial conditional logit models were estimated. The specification of the 8-choice model is unique in the grouping of explanatory variables to isolate effects of the price, food safety information, and household characteristics into commodity-specific effects. Interaction terms were included to investigate any effects from food safety information that are specific to certain groups of households and may differ from the average effect across the entire population of households. The results of the 8-choice model suggest that the individuals most likely to stop purchasing beef in a given month, when the amount of food safety information increases, are households with college educated heads. This is also the effect for households in urban areas, with respect to poultry purchases. Other types of households do not appear to have a measurable response to food safety information with regard to discrete purchase decisions of beef, pork, and poultry.

Discrete choice models differ from marginal demand models in that it is not the quantity of meat and poultry purchased that is modeled, but instead the decision to purchase. It seems plausible that consumers may respond to a food safety announcement by choosing not to buy the commodity associated with the announcement. Results of the models estimated in this chapter provide evidence, although small in magnitude, of this avoidance behavior across heterogeneous groups of households. However, avoidance behavior could also be measured as continuous rather than discrete changes in purchased quantities. Therefore, the model specifications used in this chapter are carried forward into the next chapter where analysis is conducted to address whether or not food safety information affects the quantity of meat and poultry households purchase.

Table 4.1 Summary Statistics of the Binary Choice Model Variables

	Average	Minimum	Maximum	Std. Dev.
Beef Price	3.209	0.577	12.638	0.562
Pork Price	2.534	0.627	12.219	0.509
Poultry Price	1.924	0.700	8.195	0.248
Beef MI	7.633	0.786	77.645	6.428
Pork MI	2.547	0	16.567	1.988
Poultry MI	11.378	2.000	38.310	6.054
Ed	0.393	0	1	0.488
Age	0.372	0	1	0.483
Urban	0.875	0	1	0.330
Child	0.296	0	1	0.456
Income	5.383	0.250	12.500	3.151
Income^2	38.910	0.062	156.250	43.477
Hsize	2.532	1	9	1.379
Y1	0.120	0	1	0.325
Y2	0.112	0	1	0.316
Y3	0.118	0	1	0.322
Y4	0.127	0	1	0.333
Y5	0.133	0	1	0.340
Y6	0.136	0	1	0.342
Y7	0.129	0	1	0.336
Y8	0.125	0	1	0.330
M1	0.083	0	1	0.276
M2	0.083	0	1	0.276
M3	0.083	0	1	0.276
M4	0.083	0	1	0.276
M5	0.083	0	1	0.276
M6	0.083	0	1	0.276
M7	0.083	0	1	0.276
M8	0.083	0	1	0.276
M9	0.083	0	1	0.276
M10	0.083	0	1	0.276
M11	0.083	0	1	0.276
M12	0.083	0	1	0.276
South	0.366	0	1	0.482
Central	0.204	0	1	0.403
West	0.217	0	1	0.412
Northeast	0.213	0	1	0.410
Caucasian	0.766	0	1	0.423
Hispanic	0.076	0	1	0.264
Black	0.121	0	1	0.326
Asian	0.022	0	1	0.146
Other	0.016	0	1	0.126

Note: The number of observations is 745,632.

Table 4.2 Estimated Coefficients of Binary Choice Models

	Beef		Pork		Poultry	
	Coefficient	Robust Standard Error	Coefficient	Robust Standard Error	Coefficient	Robust Standard Error
Beef Price	-1.147*	0.033	-0.147*	0.013	-0.192*	0.012
Pork Price	-0.060*	0.012	-0.776*	0.028	-0.024*	0.011
Poultry Price	-0.275*	0.024	-0.231*	0.022	-2.239*	0.061
Beef MI	-0.001	0.002	0.000	0.001	-0.001	0.001
Pork MI	0.011*	0.002	0.018*	0.006	0.013*	0.002
Poultry MI	-0.003*	0.001	-0.003*	0.001	0.001	0.002
Ed*MI_{beef}	-0.002	0.001	--	--	--	--
Age*MI_{beef}	-0.002	0.001	--	--	--	--
Child*MI_{beef}	-0.003*	0.001	--	--	--	--
Urban*MI_{beef}	0.000	0.001	--	--	--	--
Ed*MI_{pork}	--	--	-0.006	0.004	--	--
Age*MI_{pork}	--	--	-0.010*	0.004	--	--
Child*MI_{pork}	--	--	-0.009*	0.004	--	--
Urban*MI_{pork}	--	--	-0.007	0.005	--	--
Ed*$MI_{poultry}$	--	--	--	--	-0.002	0.001
Age*$MI_{poultry}$	--	--	--	--	0.002	0.001
Child*$MI_{poultry}$	--	--	--	--	-0.003	0.002
Urban*$MI_{poultry}$	--	--	--	--	-0.003	0.002
Ed	-0.023	0.041	0.022	0.037	0.011	0.038
Age	0.005	0.035	0.023	0.032	-0.024	0.033
Child	-0.005	0.034	-0.022	0.031	0.001	0.033
Urban	0.072	0.042	0.104*	0.039	0.194*	0.044
Income	0.161*	0.015	0.126*	0.014	0.127*	0.013
$Income^2$	-0.005*	0.001	-0.004*	0.001	-0.002*	0.001
Hsize	-0.031*	0.012	0.017	0.011	-0.037*	0.011
M1	0.079*	0.018	-0.256*	0.017	0.392*	0.015
M2	0.032*	0.016	-0.342*	0.016	0.293*	0.014
M3	0.133*	0.018	-0.205*	0.016	0.315*	0.015
M4	0.025	0.016	-0.131*	0.015	0.249*	0.014
M5	0.244*	0.015	-0.333*	0.016	0.386*	0.014
M6	0.074*	0.016	-0.440*	0.016	0.261*	0.014
M7	0.084*	0.016	-0.386*	0.016	0.315*	0.015
M8	0.091*	0.015	-0.343*	0.016	0.360*	0.015
M9	0.037*	0.015	-0.330*	0.016	0.270*	0.014
M10	0.064*	0.015	-0.310*	0.015	0.225*	0.014
M11	-0.160*	0.015	-0.284*	0.016	0.200*	0.018

Note: A * denotes coefficients that are statitically significantly different from zero at the 5 percent level or better.

Table 4.2 Estimated Coefficients of Binary Choice Models, cont.

	Beef		Pork		Poultry	
	Coefficient	Robust Standard Error	Coefficient	Robust Standard Error	Coefficient	Robust Standard Error
Y1	0.243*	0.030	1.195*	0.035	0.696*	0.026
Y2	-0.155*	0.029	0.227*	0.022	0.095*	0.021
Y3	-0.024	0.023	0.125*	0.019	0.180*	0.018
Y4	-0.038*	0.019	0.100*	0.017	0.085*	0.016
Y5	-0.102*	0.016	-0.035*	0.014	-0.034*	0.013
Y7	0.165*	0.016	0.071*	0.015	0.156*	0.015
Y8	-0.029	0.019	-0.001	0.018	0.174*	0.019
Central	0.130	0.149	0.118	0.125	0.087	0.113
West	0.488*	0.161	0.232*	0.119	0.387*	0.111
Northeast	0.179	0.135	0.045	0.121	0.326*	0.110
Hispanic	-0.017	0.074	0.074	0.073	0.054	0.065
Black	0.020	0.101	0.063	0.095	0.031	0.081
Asian	-0.145	0.109	-0.159*	0.080	-0.060	0.103
Other	-0.030	0.071	0.085	0.067	-0.005	0.059
Log Pseudolikelihood	-328,198.560		-323,098.210		-361,624.710	
Number of Obs	705,639		681,356		722,533	

Note: A * denotes coefficients that are statitically significantly different from zero at the 5 percent level or better.

Table 4.3 Binary Choice Model Elasticities

	Elasticity	Standard Deviation	95% Confidence Interval	
BEEF MODEL				
Price - Own	-3.417	0.099	-3.222	-3.611
Cross Price - Pork	-0.145	0.028	-0.089	-0.201
Cross Price - Poultry	-0.487	0.044	-0.401	-0.574
Income	0.465	0.028	0.520	0.409
Food Safety - Ed	-0.020	0.012	0.004	-0.044
Food Safety - Age	-0.020	0.011	0.001	-0.042
Food Safety - Child	-0.025	0.013	-0.001	-0.050
Food Safety - Urban	-0.005	0.008	0.011	-0.021
PORK MODEL				
Cross Price - Beef	-0.439	0.036	-0.368	-0.510
Price - Own	-1.870	0.071	-1.731	-2.009
Cross Price - Poultry	-0.410	0.040	-0.332	-0.488
Income	0.335	0.023	0.381	0.289
Food Safety - Ed	0.031	0.014	0.059	0.002
Food Safety - Age	0.019	0.013	0.044	-0.007
Food Safety - Child	0.022	0.015	0.050	-0.007
Food Safety - Urban	0.029	0.009	0.047	0.011
POULTRY MODEL				
Cross Price - Beef	-0.572	0.036	-0.501	-0.642
Cross Price - Pork	-0.058	0.026	-0.008	-0.109
Price - Own	-3.969	0.109	-3.755	-4.182
Income	0.506	0.025	0.555	0.457
Food Safety - Ed	-0.011	0.026	0.041	-0.062
Food Safety - Age	0.027	0.025	0.075	-0.021
Food Safety - Child	-0.020	0.026	0.031	-0.072
Food Safety - Urban	-0.025	0.015	0.004	-0.055

Table 4.4 Summary Statistics of 8-Choice Model Variables

	Full Sample				Random Sample 1			
	Average	Minimum	Maximum	Std. Dev.	Average	Minimum	Maximum	Std. Dev.
Beef Price	3.046	0.170	8.006	0.493	3.034	0.346	7.529	0.491
Pork Price	2.480	0.055	10.795	0.476	2.473	0.055	10.188	0.479
Poultry Price	1.822	0.150	6.045	0.245	1.815	0.156	4.296	0.243
Beef MI	7.633	0.786	77.645	6.428	7.650	0.786	77.645	6.446
Pork MI	2.547	0.000	16.567	1.988	2.558	0.000	16.567	2.010
Poultry MI	11.378	2.000	38.310	6.054	11.336	2.000	38.310	6.021
Ed	0.393	0	1	0.488	0.376	0	1	0.484
Age	0.372	0	1	0.483	0.376	0	1	0.484
Child	0.296	0	1	0.456	0.288	0	1	0.453
Urban	0.875	0	1	0.330	0.873	0	1	0.333
Income	5.383	0.250	12.500	3.151	5.281	0.250	12.500	3.137
Income^2	38.910	0.062	156.250	43.477	37.729	0.062	156.250	43.064
Hsize	2.532	1	9	1.379	2.527	1	9	1.359
Y1	0.120	0	1	0.325	0.120	0	1	0.325
Y2	0.112	0	1	0.316	0.114	0	1	0.318
Y3	0.118	0	1	0.322	0.118	0	1	0.323
Y4	0.127	0	1	0.333	0.130	0	1	0.337
Y5	0.133	0	1	0.340	0.131	0	1	0.338
Y6	0.136	0	1	0.342	0.134	0	1	0.341
Y7	0.129	0	1	0.336	0.130	0	1	0.336
Y8	0.125	0	1	0.330	0.122	0	1	0.328
M1	0.083	0	1	0.276	0.083	0	1	0.276
M2	0.083	0	1	0.276	0.083	0	1	0.276
M3	0.083	0	1	0.276	0.083	0	1	0.276
M4	0.083	0	1	0.276	0.083	0	1	0.276
M5	0.083	0	1	0.276	0.083	0	1	0.276
M6	0.083	0	1	0.276	0.083	0	1	0.276
M7	0.083	0	1	0.276	0.083	0	1	0.276
M8	0.083	0	1	0.276	0.083	0	1	0.276
M9	0.083	0	1	0.276	0.083	0	1	0.276
M10	0.083	0	1	0.276	0.083	0	1	0.276
M11	0.083	0	1	0.276	0.083	0	1	0.276
M12	0.083	0	1	0.276	0.083	0	1	0.276
South	0.366	0	1	0.482	0.362	0	1	0.481
Central	0.204	0	1	0.403	0.216	0	1	0.412
West	0.217	0	1	0.412	0.216	0	1	0.412
Northeast	0.213	0	1	0.410	0.205	0	1	0.404
Caucasian	0.766	0	1	0.423	0.758	0	1	0.429
Hispanic	0.076	0	1	0.264	0.075	0	1	0.264
Black	0.121	0	1	0.326	0.123	0	1	0.328
Asian	0.022	0	1	0.146	0.026	0	1	0.159
Other	0.016	0	1	0.126	0.018	0	1	0.134
# of Obs	745,632				119,280			

Table 4.4 Summary Statistics of 8-Choice Model Variables, cont.

	Random Sample 2				Random Sample 3			
	Average	Minimum	Maximum	Std. Dev.	Average	Minimum	Maximum	Std. Dev.
Beef Price	3.058	0.217	8.006	0.492	3.046	0.178	7.371	0.491
Pork Price	2.485	0.164	10.188	0.475	2.472	0.113	8.546	0.470
Poultry Price	1.828	0.156	6.045	0.244	1.817	0.223	5.068	0.244
Beef MI	7.672	0.786	77.645	6.500	7.589	0.786	77.645	6.471
Pork MI	2.563	0.000	16.567	2.010	2.513	0.000	16.567	1.964
Poultry MI	11.532	2.000	38.310	6.129	11.244	2.000	38.310	6.007
Ed	0.409	0	1	0.492	0.383	0	1	0.486
Age	0.379	0	1	0.485	0.369	0	1	0.483
Child	0.288	0	1	0.453	0.301	0	1	0.459
Urban	0.877	0	1	0.329	0.873	0	1	0.333
Income	5.447	0.250	12.500	3.187	5.305	0.250	12.500	3.100
Income^2	39.833	0.062	156.250	44.318	37.748	0.062	156.250	42.638
Hsize	2.482	1	9	1.313	2.557	1	9	1.368
Y1	0.119	0	1	0.323	0.113	0	1	0.317
Y2	0.110	0	1	0.313	0.109	0	1	0.312
Y3	0.116	0	1	0.321	0.119	0	1	0.324
Y4	0.125	0	1	0.330	0.128	0	1	0.334
Y5	0.134	0	1	0.341	0.136	0	1	0.343
Y6	0.137	0	1	0.344	0.138	0	1	0.345
Y7	0.131	0	1	0.337	0.130	0	1	0.336
Y8	0.128	0	1	0.334	0.127	0	1	0.333
M1	0.083	0	1	0.276	0.083	0	1	0.276
M2	0.083	0	1	0.276	0.083	0	1	0.276
M3	0.083	0	1	0.276	0.083	0	1	0.276
M4	0.083	0	1	0.276	0.083	0	1	0.276
M5	0.083	0	1	0.276	0.083	0	1	0.276
M6	0.083	0	1	0.276	0.083	0	1	0.276
M7	0.083	0	1	0.276	0.083	0	1	0.276
M8	0.083	0	1	0.276	0.083	0	1	0.276
M9	0.083	0	1	0.276	0.083	0	1	0.276
M10	0.083	0	1	0.276	0.083	0	1	0.276
M11	0.083	0	1	0.276	0.083	0	1	0.276
M12	0.083	0	1	0.276	0.083	0	1	0.276
South	0.390	0	1	0.488	0.342	0	1	0.474
Central	0.186	0	1	0.389	0.202	0	1	0.401
West	0.220	0	1	0.414	0.222	0	1	0.416
Northeast	0.204	0	1	0.403	0.234	0	1	0.423
Caucasian	0.771	0	1	0.420	0.770	0	1	0.421
Hispanic	0.070	0	1	0.255	0.076	0	1	0.264
Black	0.120	0	1	0.325	0.116	0	1	0.320
Asian	0.024	0	1	0.152	0.024	0	1	0.154
Other	0.016	0	1	0.125	0.014	0	1	0.117
# of Obs	745,632				119,280			

Table 4.5 Estimated Coefficients of 8-Choice Models

		Sample 1		Sample 2		Sample 3	
	Alternative	Coefficent	Robust Std Error	Coefficent	Robust Std Error	Coefficent	Robust Std Error
Price - Own	Beef	-0.977*	0.078	-0.806*	0.081	-0.828*	0.072
	Pork	-0.848*	0.087	-0.895*	0.081	-0.838*	0.090
	Poultry	-2.492*	0.166	-2.202*	0.169	-2.206*	0.162
Price - Beef	Pork	-0.016	0.049	-0.085	0.052	-0.105*	0.048
	Poultry	0.048	0.043	-0.007	0.044	0.002	0.039
Price - Pork	Beef	0.131*	0.035	0.146*	0.029	0.083*	0.035
	Poultry	0.086*	0.034	0.040	0.031	0.046	0.037
Price - Poultry	Beef	0.083	0.072	0.028	0.073	0.075	0.073
	Pork	-0.080	0.077	-0.159*	0.076	-0.131	0.076
MI - Own	Beef	-0.002	0.004	-0.001	0.003	0.002	0.004
	Pork	-0.017	0.015	-0.009	0.016	-0.018	0.016
	Poultry	-0.009	0.008	-0.001	0.008	-0.009	0.008
MI - Beef	Pork	-0.001	0.001	0.001	0.001	0.002	0.001
	Poultry	0.001	0.001	0.000	0.001	-0.003*	0.001
MI - Pork	Beef	0.008	0.005	0.004	0.004	0.002	0.005
	Poultry	0.006	0.005	0.002	0.004	0.015*	0.005
MI - Poultry	Beef	-0.002	0.002	-0.003*	0.002	-0.001	0.002
	Pork	-0.001	0.002	-0.002	0.002	-0.003*	0.002
Ed*MI	Beef	-0.007*	0.002	-0.004	0.002	-0.004	0.002
	Pork	-0.009	0.011	0.000	0.011	0.002	0.011
	Poultry	0.009	0.005	0.006	0.005	0.008	0.004
Age*MI	Beef	-0.001	0.002	-0.004	0.002	-0.005*	0.002
	Pork	0.012	0.012	0.031*	0.013	0.009	0.012
	Poultry	0.007	0.005	-0.005	0.005	0.001	0.005
Child*MI	Beef	0.002	0.003	0.000	0.003	0.001	0.003
	Pork	-0.002	0.012	0.015	0.013	-0.004	0.012
	Poultry	-0.001	0.005	-0.004	0.005	0.002	0.005
Urban*MI	Beef	0.003	0.003	0.004	0.003	-0.001	0.003
	Pork	0.016	0.014	-0.006	0.014	0.018	0.015
	Poultry	-0.002	0.007	0.000	0.007	0.005	0.007
Ed	Beef	-0.197*	0.054	-0.174*	0.053	-0.115*	0.051
	Pork	-0.184*	0.056	-0.205*	0.056	-0.279*	0.053
	Poultry	-0.131	0.069	-0.114	0.069	-0.054	0.067
Age	Beef	0.124*	0.054	0.123*	0.055	0.180*	0.054
	Pork	0.267*	0.054	0.194*	0.058	0.259*	0.055
	Poultry	-0.071	0.068	0.029	0.069	-0.016	0.069
Child	Beef	-0.144*	0.067	-0.176*	0.071	0.023	0.063
	Pork	-0.286*	0.069	-0.247*	0.068	-0.003	0.063
	Poultry	-0.035	0.080	0.019	0.081	-0.054	0.079
Urban	Beef	0.016	0.077	-0.053	0.076	0.143	0.080
	Pork	-0.096	0.073	0.020	0.078	-0.078	0.076
	Poultry	0.251*	0.098	0.133	0.097	0.089	0.097
Income	Beef	0.149*	0.028	0.170*	0.028	0.111*	0.027
	Pork	0.111*	0.028	0.129*	0.026	0.163*	0.026
	Poultry	0.121*	0.025	0.170*	0.025	0.157*	0.025
Income^2	Beef	-0.006*	0.002	-0.007*	0.002	-0.004	0.002
	Pork	-0.005*	0.002	-0.004*	0.002	-0.007*	0.002
	Poultry	-0.001	0.002	-0.004	0.002	-0.004*	0.002
Hsize	Beef	0.064*	0.025	0.092*	0.026	0.008	0.024
	Pork	0.075*	0.023	0.056*	0.022	0.017	0.021
	Poultry	0.021	0.021	0.042	0.023	0.046*	0.021

Note: A * denotes coefficients that are statitically significantly different from zero at the 5 percent level or better.

Table 4.5 Estimated Coefficients of 8-Choice Models, cont.

		Sample 1		Sample 2		Sample 3	
	Alternative	Coefficent	Robust Std Error	Coefficent	Robust Std Error	Coefficent	Robust Std Error
Y1	Beef	-0.375*	0.074	-0.214*	0.072	-0.188*	0.070
	Pork	1.048*	0.098	0.973*	0.092	0.862*	0.098
	Poultry	0.452*	0.062	0.405*	0.059	0.390*	0.064
Y2	Beef	-0.375*	0.061	-0.213*	0.064	-0.179*	0.059
	Pork	0.117*	0.051	0.043	0.052	-0.023	0.051
	Poultry	-0.009	0.048	-0.063	0.048	-0.079	0.047
Y3	Beef	-0.192*	0.050	-0.075	0.052	-0.131*	0.048
	Pork	0.091*	0.042	0.038	0.043	-0.030	0.043
	Poultry	0.067	0.040	0.065	0.040	0.078*	0.040
Y4	Beef	-0.106*	0.041	-0.095*	0.043	-0.055	0.041
	Pork	0.156*	0.038	0.109*	0.038	0.058	0.037
	Poultry	0.006	0.035	0.038	0.036	0.047	0.036
Y5	Beef	-0.077*	0.035	-0.111*	0.035	-0.044	0.034
	Pork	0.001	0.033	-0.025	0.031	-0.083*	0.032
	Poultry	-0.087*	0.031	-0.035	0.031	-0.045	0.030
Y7	Beef	0.131*	0.036	0.149*	0.038	0.135*	0.036
	Pork	0.130*	0.036	0.159*	0.036	0.166*	0.035
	Poultry	0.194*	0.037	0.200*	0.037	0.190*	0.036
Y8	Beef	-0.009	0.041	0.013	0.040	0.019	0.041
	Pork	0.072	0.043	0.097*	0.041	0.151*	0.042
	Poultry	0.320*	0.046	0.336*	0.045	0.275*	0.046
M1	Beef	0.024	0.037	0.053	0.039	0.070	0.038
	Pork	-0.329*	0.036	-0.358*	0.036	-0.418*	0.036
	Poultry	0.301*	0.034	0.395*	0.034	0.373*	0.033
M2	Beef	0.033	0.034	0.055	0.035	0.097*	0.034
	Pork	-0.381*	0.033	-0.366*	0.033	-0.416*	0.034
	Poultry	0.303*	0.031	0.318*	0.032	0.290*	0.031
M3	Beef	0.100*	0.038	0.111*	0.040	0.149*	0.038
	Pork	-0.310*	0.035	-0.317*	0.036	-0.363*	0.035
	Poultry	0.260*	0.033	0.317*	0.033	0.255*	0.032
M4	Beef	0.002	0.033	0.017	0.034	0.004	0.034
	Pork	-0.203*	0.032	-0.183*	0.034	-0.206*	0.034
	Poultry	0.216*	0.032	0.268*	0.031	0.236*	0.032
M5	Beef	0.212*	0.033	0.187*	0.033	0.237*	0.034
	Pork	-0.373*	0.035	-0.393*	0.034	-0.417*	0.034
	Poultry	0.336*	0.032	0.410*	0.031	0.358*	0.032
M6	Beef	0.137*	0.032	0.082*	0.034	0.153*	0.034
	Pork	-0.421*	0.034	-0.430*	0.035	-0.501*	0.036
	Poultry	0.290*	0.032	0.350*	0.033	0.294*	0.032
M7	Beef	0.101*	0.035	0.099*	0.035	0.131*	0.035
	Pork	-0.375*	0.037	-0.368*	0.037	-0.466*	0.037
	Poultry	0.302*	0.034	0.360*	0.033	0.380*	0.034
M8	Beef	0.114*	0.033	0.086*	0.033	0.131*	0.033
	Pork	-0.385*	0.036	-0.320*	0.035	-0.366*	0.036
	Poultry	0.383*	0.035	0.459*	0.034	0.383*	0.034
M9	Beef	0.056	0.033	0.061	0.033	0.116*	0.034
	Pork	-0.347*	0.034	-0.291*	0.034	-0.388*	0.035
	Poultry	0.298*	0.033	0.324*	0.032	0.291*	0.032
M10	Beef	0.077*	0.033	0.076*	0.034	0.103*	0.033
	Pork	-0.346*	0.034	-0.340*	0.034	-0.378*	0.033
	Poultry	0.224*	0.031	0.284*	0.032	0.249*	0.031
M11	Beef	-0.156*	0.032	-0.189*	0.032	-0.134*	0.033
	Pork	-0.287*	0.036	-0.295*	0.037	-0.339*	0.036
	Poultry	0.081	0.042	0.262*	0.042	0.215*	0.041

Note: A * denotes coefficients that are statitically significantly different from zero at the 5 percent level or better.

Table 4.5 Estimated Coefficients of 8-Choice Models, cont.

		Sample 1		Sample 2		Sample 3	
	Alternative	Coefficent	Robust Std Error	Coefficent	Robust Std Error	Coefficent	Robust Std Error
Central	Beef	-0.105	0.070	-0.074	0.069	-0.058	0.071
	Pork	0.072	0.063	0.140*	0.062	0.108	0.063
	Poultry	-0.226*	0.061	-0.208*	0.063	-0.212*	0.062
West	Beef	0.053	0.067	0.078	0.070	0.053	0.067
	Pork	-0.140*	0.065	0.048	0.063	-0.129*	0.061
	Poultry	0.343*	0.060	0.278*	0.060	0.347*	0.058
Northeast	Beef	-0.013	0.063	0.120	0.062	0.069	0.060
	Pork	-0.041	0.059	0.116	0.059	0.059	0.057
	Poultry	0.143*	0.055	0.238*	0.056	0.195*	0.056
Hispanic	Beef	-0.136	0.077	-0.077	0.080	0.068	0.080
	Pork	-0.083	0.081	-0.110	0.078	-0.100	0.077
	Poultry	0.206*	0.067	0.155*	0.074	0.018	0.074
Black	Beef	-0.622*	0.063	-0.547*	0.064	-0.715*	0.065
	Pork	0.159*	0.068	0.144*	0.068	0.123	0.068
	Poultry	0.539*	0.058	0.574*	0.058	0.477*	0.059
Asian	Beef	-0.693*	0.160	-0.180	0.128	-0.137	0.159
	Pork	0.383*	0.153	0.368*	0.155	0.228	0.155
	Poultry	0.209	0.111	0.141	0.137	0.070	0.128
Other	Beef	-0.474*	0.142	-0.444*	0.217	-0.545*	0.152
	Pork	0.247*	0.123	-0.025	0.200	-0.032	0.139
	Poultry	0.005	0.131	0.155	0.134	0.119	0.133
Constant	Alternative 1	1.450*	0.280	0.873*	0.277	1.051*	0.274
	Alternative 2	-0.031	0.298	0.278	0.289	0.330	0.301
	Alternative 3	1.892*	0.333	1.422*	0.331	1.607*	0.324
	Alternative 4	2.916*	0.435	2.589*	0.429	2.807*	0.436
	Alternative 5	4.431*	0.476	3.353*	0.467	3.687*	0.431
	Alternative 6	2.931*	0.509	2.697*	0.464	2.918*	0.484
	Alternative 7	6.554*	0.633	5.752*	0.590	6.108*	0.589
Log Pseudo-likelihood		-216,716.950		-217,282.180		-215,293.090	

Note: A * denotes coefficients that are statitically significantly different from zero at the 5 percent level or better.

Table 4.6 8-Choice Model Elasticities

	Commodity	Elasticity	Standard Deviation	95% Confidence Interval	
Own Price	Beef	-1.309	0.106	-1.102	-1.517
	Pork	-1.376	0.139	-1.104	-1.647
	Poultry	-2.436	0.171	-2.101	-2.771
Income	Beef	0.140	0.021	0.181	0.099
	Pork	0.152	0.033	0.217	0.087
	Poultry	0.293	0.027	0.346	0.241
Food Safety - Ed	Beef	-0.033	0.015	-0.004	-0.062
	Pork	-0.049	0.030	0.010	-0.108
	Poultry	-0.002	0.050	0.095	-0.099
Food Safety - Age	Beef	-0.010	0.011	0.011	-0.031
	Pork	-0.008	0.023	0.036	-0.052
	Poultry	-0.013	0.042	0.070	-0.095
Food Safety - Child	Beef	0.000	0.012	0.024	-0.023
	Pork	-0.034	0.028	0.020	-0.088
	Poultry	-0.058	0.043	0.026	-0.143
Food Safety - Urban	Beef	0.003	0.007	0.017	-0.012
	Pork	0.000	0.016	0.031	-0.032
	Poultry	-0.072	0.024	-0.025	-0.119

5 Chapter

Demand Models of Meat and Poultry Purchases

5.1 Introduction

The final models employed in this analysis of food safety information on monthly household purchases of meat and poultry is the estimation of a continuous demand system. Unlike the discrete choice models of the previous chapter, the continuous models used in this chapter are intended to capture changes in the quantity of meat and poultry purchases due to changes in the amount of food safety information available. The use of a system estimator will allow for any correlation that exists between disturbances of the demand equations for beef, pork, and poultry.

The discrete choice framework of the previous chapter was employed based on the assumption that avoidance of meat or poultry in the wake of a food safety event is rational consumer behavior. The idea of an avoidance response does not have to be abandoned in order to employ a marginal model at the household level. One possible reason for a measurable marginal response of avoidance behavior is that recalls are product specific. That is, a recall may include ground beef, for example, but beef roasts and steaks are not mentioned. The commodity-level data used in this analysis is aggregated over all fresh beef products, so avoidance of one beef product (e.g. ground beef) does not imply avoidance of all beef products. This response would be measured as a decline in total beef purchases, rather than a complete avoidance of all beef products, as measured by a discrete model.

The objective of the analysis presented in this chapter is to determine if food safety information available to the public impacts the amount of meat and poultry purchased by heterogeneous households. Two different system demand models are estimated: a cross-sectional (pooled) model and a panel (random-effects) model. The estimation results of the two models indicate that accounting for the panel aspect of the Nielsen Homescan data has a distinct difference on the conclusion that can be made about food safety impacts on household purchases.

5.2 Demand System Model

The demand model specified in this chapter 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 proportion of censoring found in the monthly household purchase data ranges from 42% to 65%, depending on the commodity. This level of censoring necessitates the use of a tobit estimator to account for both zero and positive meat and poultry purchases in an unbiased manner. 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 models. 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. Two different versions of the SUR tobit model are presented in this chapter. The first is a pooled SUR tobit model where the observations are considered to be purely cross-sectional and the panel aspect of the household data is not accounted for in the model. The second version of the SUR tobit model is a random effects estimator, which explicitly accounts for households that appear more than once in the data through a component error structure. These two models are presented in the following sections.

5.2.1 Pooled SUR Tobit

Consider a SUR tobit model with J commodities (equations) and N outcomes. A single household may provide multiple outcomes, but the panel nature of the data is not explicitly accounted for in this model. The observed dependent variable y_{ij} is determined by:

$$y_{ij}^* = \alpha_j + \mathbf{x}_{ij}\boldsymbol{\beta}_j + \mathbf{c}_i\boldsymbol{\gamma}_j + \varepsilon_{ij} \ , \quad j = 1,\ldots,J, \ i = 1,\ldots,N \ , \tag{5.1}$$

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{ij}^* > 0 \\ 0 & \text{if } y_{ij}^* \le 0 \end{cases} , \tag{5.2}$$

where y_{ij}^{*} is the latent or unobserved dependent variable, α_j is the intercept term for the j^{th} commodity, $\mathbf{x}_{ij}$ is a $(1\times k_x)$ vector of commodity-specific regressors, $\boldsymbol{\beta}_j$ is the corresponding $(k_x\times 1)$ vector of unknown coefficients, $\mathbf{c}_i$ is a $(1\times k_c)$ vector of household-specific regressors, $\boldsymbol{\gamma}_j$ is the corresponding $(k_c\times 1)$ vector of unknown coefficients, and $\boldsymbol{\varepsilon}_i=[\varepsilon_{i1},...,\varepsilon_{iJ}]\sim iid\ N(0,\Sigma)$ where Σ is a $(J\times J)$ symmetric positive definite matrix.[18] According to equation (5.2), the value of the observed dependent variable, y_{ij}, is equal to the latent value if y_{ij}^{*} is greater than zero and censored at zero otherwise.

The equations in the system can be stacked over commodities and rewritten as:

$$\begin{bmatrix} y_{i1}^{*} \\ y_{i2}^{*} \\ \vdots \\ y_{iJ}^{*} \end{bmatrix}=\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_J \end{bmatrix}+\begin{bmatrix} \mathbf{x}_{i1} & 0 & \cdots & 0 \\ 0 & \mathbf{x}_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_{iJ} \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_J \end{bmatrix}+\begin{bmatrix} \mathbf{c}_i & 0 & \cdots & 0 \\ 0 & \mathbf{c}_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{c}_i \end{bmatrix}\begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\gamma}_J \end{bmatrix}+\begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iJ} \end{bmatrix},\ i=1,...,N, \quad (5.3)$$

or

$$\mathbf{y}_i^{*}=\boldsymbol{\alpha}+X_{\mathbf{i}}\boldsymbol{\beta}+C_{\mathbf{i}}\boldsymbol{\gamma}+\boldsymbol{\varepsilon}_i\ . \quad (5.4)$$

Combining the regressor matrices, X_i and C_i, equation (5.4) can be rewritten as:

$$\mathbf{y}_i^{*}=W_i\theta+\boldsymbol{\varepsilon}_i\ , \quad (5.5)$$

where $W_i=\left[I_j\ X_i\ C_i\right]$, $\theta=[\boldsymbol{\alpha}\ \boldsymbol{\beta}\ \boldsymbol{\gamma}]$, and I_J is a $(J\times J)$ identity matrix. The system equations are further stacked over all households in the dataset and written as:

[18] If $\mathrm{E}(\varepsilon\varepsilon')=\sigma_{\varepsilon_j}^{2}I_J$ where I_J is a $(J\times J)$ identity matrix, then the SUR estimator is equivalent to single equation Tobit estimation of the individual demand equations.

$$\begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \\ \vdots \\ \mathbf{y}_N^* \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_N \end{bmatrix} \theta + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix} \tag{5.6}$$

or

$$\mathbf{y}^* = W \cdot \theta + \boldsymbol{\varepsilon} , \tag{5.7}$$

where $\mathbf{y}^*$ is $(N \cdot J \times 1)$, W is $(N \cdot J \times K)$, θ is $(K \times 1)$, $\boldsymbol{\varepsilon}$ is $(N \cdot J \times 1)$, and $K = J + k_x + k_c$ is the total number of parameters to be estimated.

5.2.2 Random Effects SUR Tobit

The random effects SUR tobit model is specified in a similar manner to the pooled SUR tobit model given in equation (5.1). The difference in the two models arises from the specification of a component error structure 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:

$$y_{ijt}^* = \alpha_j + \mathbf{x}_{ijt}\boldsymbol{\beta}_j + \mathbf{c}_i\boldsymbol{\gamma}_j + u_{ij} + \varepsilon_{ijt} \ , \quad j = 1,...,J,\ i = 1,...,N,\ t = 1,...,T_i \ , \tag{5.8}$$

$$y_{ijt} = \begin{cases} y_{ijt}^* & \text{if } y_{ijt}^* > 0 \\ 0 & \text{if } y_{ijt}^* \le 0 \end{cases} , \tag{5.9}$$

where u_{ij} is the household- and commodity-specific random error term that does not vary over time, $u_{ij} \sim iid\ N\left(0, \sigma_{u_j}^2\right)$, T_i is the size of the panel for the i^{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:

$$\begin{bmatrix} y_{it1}^* \\ y_{it2}^* \\ \vdots \\ y_{itJ}^* \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_J \end{bmatrix} + \begin{bmatrix} \mathbf{x}_{it1} & 0 & \cdots & 0 \\ 0 & \mathbf{x}_{it2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_{itJ} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_J \end{bmatrix} + \begin{bmatrix} \mathbf{c}_i & 0 & \cdots & 0 \\ 0 & \mathbf{c}_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{c}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\gamma}_J \end{bmatrix} + \begin{bmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{iJ} \end{bmatrix} + \begin{bmatrix} \varepsilon_{it1} \\ \varepsilon_{it2} \\ \vdots \\ \varepsilon_{itJ} \end{bmatrix}, \quad (5.10)$$

or

$$\mathbf{y}_{it}^* = \boldsymbol{\alpha} + X_{it}\boldsymbol{\beta} + C_i\boldsymbol{\gamma} + \mathbf{u}_i + \boldsymbol{\varepsilon}_{it} \ , \quad (5.11)$$

for $i = 1, ..., N$, $t = 1, ..., T_i$. Combining the regressor matrices, X_{it} and C_i, equation (5.11) can be rewritten as:

$$\mathbf{y}_{it}^* = W_{it}\theta + \mathbf{u}_i + \boldsymbol{\varepsilon}_{it} \ , \quad (5.12)$$

where $W_{it} = \left[I_j \ X_{it} \ C_i \right]$, $\theta = \left[\boldsymbol{\alpha} \, \boldsymbol{\beta} \, \boldsymbol{\gamma} \right]$, I_J is a $\left(J \times J\right)$ identity matrix, $\mathbf{u}_i \sim iid\ N\left(0, V\right)$, and $\boldsymbol{\varepsilon}_{it} \sim iid\ N\left(0, \Sigma\right)$ with $\mathrm{E}\left(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{is}'\right) = 0$ for all $t \neq s$. The covariance matrix V is defined as follows:

$$V = \begin{bmatrix} \sigma_{u_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{u_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{u_J}^2 \end{bmatrix}. \tag{5.13}$$

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

$$\begin{bmatrix} \mathbf{y}_{11}^* \\ \vdots \\ \mathbf{y}_{1T_1}^* \\ \mathbf{y}_{21}^* \\ \vdots \\ \mathbf{y}_{2T_2}^* \\ \vdots \\ \mathbf{y}_{N1}^* \\ \vdots \\ \mathbf{y}_{NT_N}^* \end{bmatrix} = \begin{bmatrix} W_{11} \\ \vdots \\ W_{1T_1} \\ W_{21} \\ \vdots \\ W_{2T_2} \\ \vdots \\ W_{N1} \\ \vdots \\ W_{NT_N} \end{bmatrix} \theta + \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_N \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{11} \\ \vdots \\ \boldsymbol{\varepsilon}_{1T_1} \\ \boldsymbol{\varepsilon}_{21} \\ \vdots \\ \boldsymbol{\varepsilon}_{2T_2} \\ \vdots \\ \boldsymbol{\varepsilon}_{N1} \\ \vdots \\ \boldsymbol{\varepsilon}_{NT_N} \end{bmatrix}, \tag{5.14}$$

or

$$\mathbf{y}^* = W \cdot \theta + \mathbf{u} + \boldsymbol{\varepsilon}, \tag{5.15}$$

where $\mathbf{y}^*$ is $(N \cdot J \cdot T \times 1)$, W is $(N \cdot J \cdot T \times K)$, θ is $(K \times 1)$, $\mathbf{u}$ is $(N \cdot J \cdot T \times 1)$ with the same value for the i^{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.

5.3 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:

$$S = \left[S_1 = (0,...,0)', ..., S_h = \left(\underbrace{0,...,0,}_{r} \underbrace{+,...,+}_{p-r} \right)', ..., S_{2^p} = (+,...,+)' \right], \tag{5.16}$$

where S_k is $(p \times 1)$, $k = 1, 2, ..., 2^p$, 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 pooled SUR tobit model. The likelihood function for the i^{th} household in the S_h case is given by:

$$L_i^{S_h}\left(y_i \middle| W, \Sigma\right) = \int_{-\infty}^{-W_{1i}'\theta_1} ... \int_{-\infty}^{-W_{ri}'\theta_r} \left\{ (2\pi)^{-p/2} \left|\Sigma^{-1}\right|^{1/2} \exp - \tfrac{1}{2}\left(y_i^* - W_i\theta\right)' \Sigma^{-1} \left(y_i^* - W_i\theta\right) \right\} . \tag{5.17}$$

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.[19]

[19] 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

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. Huang, Sloan, and Adamache (1987) proposed a data augmentation methodology where estimation is done via an expectation-maximization (EM) algorithm. However, their methodology was limited to a bivariate SUR tobit model. An additional limitation of the EM algorithm methodology is that an estimate of the information matrix is not automatically generated. Therefore, a bootstrapping technique would have to be implemented to provide some estimate of the covariance matrix. Given the size of the dataset used in this study, this could be a very inefficient process. Therefore, a data augmentation method that does not require the use of bootstrapping to obtain estimates of all the parameters of interest is preferable.

The EM algorithm is an estimation technique that is employed in a classical framework via maximization of a likelihood function. Alternatively, the implementation of a Bayesian analysis 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

(2001)). The techniques used in these studies vary widely, suggesting that a general consensus on estimation methodology does not exist.

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).

The inference objective in Bayesian analysis is to characterize the uncertainty of any true value, such as model parameters, with a probability distribution. This distribution can then be updated with current data to get a posterior probability distribution for a parameter that has less uncertainty (Lynch, pp.50-51, 2007). The Gibbs sampler is widely used in Bayesian analysis because it can be employed in cases where sampling from the multivariate posterior distribution is not possible, but sampling from the conditional distribution for each parameter is feasible.

To illustrate the technique, suppose the following joint density $f\left(x_1, x_2\right)$ with conditional densities of $f\left(x_1 \middle| x_2\right)$ and $f\left(x_2 \middle| x_1\right)$. The Gibbs sampler draws iteratively from the conditional distributions using the most recent draw as the conditioning value. Given a starting value of x_1^{t-1}, a draw is made from the conditional distribution $f\left(x_2^t \middle| x_1^{t-1}\right)$ to obtain a value of x_2. Then using the most recent realization of x_2, a draw is taken from the conditional distribution $f\left(x_1^t \middle| x_2^t\right)$ to obtain a value of x_1. A new realization of x_2 can then be drawn from $f\left(x_2^{t+1} \middle| x_1^t\right)$ using the previous value of x_1. This process is repeated until convergence is reached, whereby the draws are from the target joint distribution $f\left(x_1, x_2\right)$.

5.3.1 Bayesian Estimation of the Pooled SUR Tobit Model

As mentioned previously, 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 pooled SUR tobit model, stacked over all J commodities is specified as:

$$\mathbf{y}_i^* = W_i\theta + \boldsymbol{\varepsilon}_i \ , \tag{5.18}$$

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{ij}^* > 0 \\ 0 & \text{if } y_{ij}^* \le 0 \ , \end{cases} \tag{5.19}$$

where $\varepsilon_i \sim iid\ N(0,\Sigma)$. The prior distributions of the unknown model parameters, $\pi(\theta)$ and $\pi(\Sigma)$, are specified as a multivariate normal and inverse Wishart distributions, respectively. The probability distribution of the dependent variable conditional on the model parameters and observed data for household i is:

$$p\left(\mathbf{y}_i \middle| \theta, W, \Sigma\right) = \int_{-\infty}^{-W_{1i}'\theta_1} \cdots \int_{-\infty}^{-W_{ri}'\theta_r} f\left(\mathbf{y}_i^* \middle| \theta, W, \Sigma\right) d\mathbf{y}_{i1}^* \ldots d\mathbf{y}_{ir}^* \ , \tag{5.20}$$

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

$$L\left(\mathbf{y}\middle|\theta,W,\Sigma\right)=\prod_{i=1}^{N}p\left(\mathbf{y}_i\middle|\theta,W,\Sigma\right)=p\left(\mathbf{y}\middle|\theta,W,\Sigma\right)\ . \tag{5.21}$$

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

$$p\left(\mathbf{y}\middle|\theta,W,\Sigma\right)\propto L\left(\mathbf{y}\middle|\theta,W,\Sigma\right)\cdot\pi\left(\theta\right)\cdot\pi\left(\Sigma\right)\ . \tag{5.22}$$

Obtaining summary statistics such as the mean, median, and variance requires integration of the probability density function. This is a difficult task because no analytical form exists for the multivariate posterior distribution given in equation (5.22). A computationally simpler method of characterizing the posterior distribution is to work with a full posterior that includes the latent variables, derive the complete posterior conditional distributions, and then use the Gibbs sampler to iteratively draw realizations of model parameters as well as the latent data. Using properties of probability distributions, the full posterior can be written as follows:

$$\begin{aligned} p\left(\theta,\Sigma,\mathbf{y}^*\middle|\mathbf{y},W\right) &\propto p\left(\mathbf{y},\mathbf{y}^*\middle|\theta,W,\Sigma\right)\cdot\pi\left(\theta\right)\cdot\pi\left(\Sigma\right) \\ &\propto p\left(\mathbf{y}\middle|\mathbf{y}^*,\theta,W,\Sigma\right)\cdot p\left(\mathbf{y}^*\middle|\mathbf{y},\theta,W,\Sigma\right)\cdot\pi\left(\theta\right)\cdot\pi\left(\Sigma\right)\ , \end{aligned} \tag{5.23}$$

which, when $\mathbf{y}^*$ is integrated out, results in the posterior expression in equation (5.22).[20]

With the full posterior in hand, the Gibbs sampler proceeds by iteratively sampling from a complete set of conditional distributions. The iterative process of sampling from the conditional posteriors is done in the following order:

[20] It can be shown that integrating the full posterior of the model parameters over the latent data will result in a posterior for those parameters that is unchanged by the addition of the latent data. Augmenting the posterior with the latent data will not alter the inference of the model parameters, but it will make the problem easier by allowing derivation of the conditional posteriors. See Koop, Poirier, and Tobias (pp.204-206, 2007) for an example of this using a probit model.

$$(1) \qquad p\left(\mathbf{y}^{*} \middle| \theta, \Sigma, \mathbf{y}, W\right) \tag{5.24}$$

$$(2) \qquad p\left(\Sigma \middle| \mathbf{z}, \theta, W\right)$$

$$(3) \qquad p\left(\theta \middle| \mathbf{z}, \Sigma, W\right),$$

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 first step of the Gibbs sampler is to augment the vector of censored purchases with draws from the conditional distribution. Let $\mathbf{z}_i = \left(\mathbf{y}^{*}_{i,r}, \mathbf{y}_{i,-r}\right)$ be a vector of dependent variables for the i^{th} household with r denoting elements censored at zero and -r denoting positive (observed) commodity purchases. The conditional distribution of $\mathbf{y}^{*}_{i,r}$ is a truncated normal distribution of the following form:

$$\mathbf{y}^{*}_{i,r} \left| \theta, \Sigma, W_i, \mathbf{y}_{i,-r} \right. \sim TN_{(-\infty,0]}\left(\boldsymbol{\mu}_{i,r}, \Sigma_r\right) \; , \tag{5.25}$$

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

$$\boldsymbol{\mu}_{i,r} = W_{i,r}\theta + \Sigma'_{r,-r}\Sigma^{-1}_{-r,-r}\left(\mathbf{y}_{i,-r} - W_{i,-r}\theta\right) \tag{5.26}$$

$$\Sigma_r = \Sigma_{r,r} + \Sigma'_{r,-r}\Sigma^{-1}_{-r,-r}\Sigma_{r,-r} \quad ,$$

where the dimension of $\boldsymbol{\mu}_{i,r}$ is $(r \times 1)$, Σ_r is dimension $(r \times r)$, and the indices r and $-r$ refer to censored and positive elements, respectively (Huang, 2001). The first step in the Gibbs sampler draws $\mathbf{z}_i$ for all people in the sample from the truncated normal distribution given in equation (5.25).

Once the first step of the Gibbs sampler is completed, $\mathbf{z}$ is a fully augmented vector that can be subsequently used for drawing realizations of the parameters of interest from the conditional distributions for the model parameters. The posterior distributions are derived from specifications of prior distributions, which convey any known information about the parameters of interest. The prior distributions for the parameters of the pooled SUR tobit model are assumed independent and of the following form:

$$\pi(\theta) \sim N_K\left(\beta_0, B_0^{-1}\right) , \tag{5.27}$$

$$\pi(\Sigma) \sim IW_J\left(\rho_0, R_0\right) , \tag{5.28}$$

where $\pi(\theta)$ is a K-dimension multivariate normal distribution with mean β_0 and precision matrix B_0^{-1} and $\pi(\Sigma)$ is a J-dimension inverse Wishart distribution with degrees of freedom ρ_0 and scale R_0. The hyperparameters of the prior distributions (β_0, B_0^{-1}, ρ_0, R_0) are assumed to be known by the researcher and are set to values that reflect any prior beliefs about the parameters θ and Σ. Theory provides little prior information on the parameters of a demand system. Therefore, the hyperparameters are set to values that reflect very diffuse prior information. The values of β_0, B_0^{-1}, ρ_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 full conditional densities of the model parameters are:

$$p\left(\theta \middle| \Sigma, W, \mathbf{z}\right) \sim N_K\left(\beta_1, B_1^{-1}\right) , \tag{5.29}$$

$$p\left(\Sigma\middle|\theta, W, \mathbf{z}\right) \sim IW_J\left(\rho_1, R_1\right). \tag{5.30}$$

The posterior distribution of θ is a K-dimension multivariate normal with mean $\beta_1 = \left(\sum_{i=1}^{N} W_i' \Sigma^{-1} W_i\right)^{-1} \left(\sum_{i=1}^{N} W_i' \Sigma^{-1} \mathbf{z}_i\right)$ and covariance matrix $B_1^{-1} = \left(\sum_{i=1}^{N} W_i' \Sigma^{-1} W_i\right)^{-1}$. The posterior distribution of Σ is a J-dimension inverse Wishart with degrees of freedom $\rho_1 = J + N$ and scale $R_1 = \left(I_J J + \bar{S}N\right) / \left(J + N\right)$, where $\bar{S} = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i - W_i\theta\right)\left(\mathbf{z}_i - W_i\theta\right)'$.

An outline of the steps of the Gibbs sampler for estimation of the pooled model is now given using the full conditional distributions derived above. The algorithm begins by assigning starting values to the parameters and the latent values of the dependent variable.[21] The Gibbs sampler algorithm is comprised of the following steps:

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

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

(2) At iteration p, complete the following:

a. Draw realizations of $\mathbf{y}_{i,r}^{*p} \middle| \theta^{p-1}, \Sigma^{p-1}, W_i, \mathbf{y}_{i,-r}$ for i=1,…,N from $TN_{(-\infty,0]}\left(\boldsymbol{\mu}_{i,r}^{p-1}, \Sigma_r^{p-1}\right)$, where $\boldsymbol{\mu}_{i,r}$ and Σ_r are person specific as described above and depend on parameters θ^{p-1} and Σ^{p-1}. The inversion method is

[21] The starting values for both the pooled and random effects SUR Tobit models are the OLS estimates of the SUR model.

used to draw from the truncated multivariate normal distribution given the most recent draws of the mean and variance of the distribution.[22]

b. Draw $\Sigma^{p} \left| \theta^{p}, \mathbf{z}^{p}, W \right.$ from $IW_{J}\left(\rho_{1}, R_{1}\right)$.

c. Draw $\theta^{p} \left| \Sigma^{p-1}, \mathbf{z}^{p}, W \right.$ from $N_{K}\left(\beta_{1}, B_{1}^{-1}\right)$.

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

5.3.2 Bayesian Estimation of the Pooled SUR Tobit Model

Estimation of the random effects SUR tobit model is conducted using the same Bayesian framework as described above for the pooled model. There are differences, however, in the conditioning elements of the posterior distributions and the number of steps in the Gibbs sampler for the random effects model. This is due to an additional unknown model parameter for the household specific error component of the model.

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

$$\mathbf{y}_{it}^{*} = W_{it}\theta + \mathbf{u}_{i} + \boldsymbol{\varepsilon}_{it} \ , \tag{5.31}$$

$$y_{ijt} = \begin{cases} y_{ijt}^{*} & \text{if } y_{ijt}^{*} > 0 \\ 0 & \text{if } y_{ijt}^{*} \leq 0 \ , \end{cases} \tag{5.32}$$

[22] The inversion method is comprised of two steps: (1) Draw a random number u from the uniform distribution $U\left(0,1\right)$. (2) Calculate $z = F^{-1}\left(u\right)$, which represents a draw from the target distribution, $f\left(x\right)$.

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, $\pi(\theta)$, 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:

$$p\left(\mathbf{y}_{it} \middle| \theta, W, \Sigma, \mathbf{u}_i\right) = \int_{-\infty}^{-W_{1it}'\theta_1} \ldots \int_{-\infty}^{-W_{rit}'\theta_r} f\left(\mathbf{y}_{it}^* \middle| \theta, W, \Sigma, \mathbf{u}_i\right) d\mathbf{y}_{it1}^* \ldots d\mathbf{y}_{itr}^* \ , \tag{5.33}$$

where $f(\bullet)$ 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:

$$L\left(\mathbf{y} \middle| \theta, W, \Sigma, \mathbf{u}\right) = \prod_{i=1}^{N} \prod_{t=1}^{T} p\left(\mathbf{y}_{it} \middle| \theta, W, \Sigma, \mathbf{u}_i\right) = p\left(\mathbf{y} \middle| \theta, W, \Sigma, \mathbf{u}\right) \ . \tag{5.34}$$

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

$$p\left(\mathbf{y} \middle| \theta, W, \Sigma, \mathbf{u}\right) \propto L\left(\mathbf{y} \middle| \theta, W, \Sigma, \mathbf{u}\right) \cdot \pi(\theta) \cdot \pi(\Sigma) \cdot \pi(V) \ . \tag{5.35}$$

As with the pooled SUR tobit model, no analytical form exists for the multivariate posterior distribution given in equation (5.35), 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:

$$\begin{aligned} p\left(\theta, \Sigma, \mathbf{u}, \mathbf{y}^* \middle| \mathbf{y}, W\right) &\propto p\left(\mathbf{y}, \mathbf{y}^* \middle| \theta, W, \Sigma, \mathbf{u}\right) \cdot \pi(\theta) \cdot \pi(\Sigma) \cdot \pi(V) \\ &\propto p\left(\mathbf{y} \middle| \mathbf{y}^*, \theta, W, \Sigma, \mathbf{u}\right) \cdot p\left(\mathbf{y}^* \middle| \theta, W, \Sigma, \mathbf{u}\right) \cdot \pi(\theta) \cdot \pi(\Sigma) \cdot \pi(V) \ . \end{aligned} \tag{5.36}$$

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:

$$\begin{aligned}
&(1) \qquad p\left(\mathbf{y}^{*} \middle| \theta, \Sigma, \mathbf{u}, W, \mathbf{y}\right) \\
&(2) \qquad p\left(V \middle| \mathbf{z}, \theta, \Sigma, \mathbf{u}, W\right) \\
&(3) \qquad p\left(\mathbf{u} \middle| \mathbf{z}, \theta, \Sigma, V, W\right) \\
&(4) \qquad p\left(\Sigma \middle| \mathbf{z}, \theta, \mathbf{u}, W\right) \\
&(5) \qquad p\left(\theta \middle| \mathbf{z}, \Sigma, \mathbf{u}, W\right),
\end{aligned} \tag{5.37}$$

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

The truncated normal distribution used in the first step of the Gibbs sample is altered slightly for the random effects SUR tobit model. The distribution must now be conditioned on the household-specific error component $\mathbf{u}_i$, which enters the mean of the distribution. Let $\mathbf{z}_{it} = \left(\mathbf{y}^*_{it,r}, \mathbf{y}_{it,-r}\right)$ be a vector of dependent variables for the i^{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:

$$\mathbf{y}^*_{it,r} \left| \theta, \Sigma, \mathbf{u}_i, W_{it}, \mathbf{y}_{it,-r} \sim TN_{(-\infty,0]}\left(\boldsymbol{\mu}_{it,r}, \Sigma_r\right) \right., \tag{5.38}$$

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^{th} household, the mean and variance of the truncated normal are:

$$\boldsymbol{\mu}_{it,r} = \mathbf{u}_i + W_{it,r}\theta + \Sigma'_{r,-r}\Sigma^{-1}_{-r,-r}\left(\mathbf{y}_{it,-r} - \mathbf{u}_i - W_{it,-r}\theta\right) \tag{5.39}$$

$$\Sigma_r = \Sigma_{r,r} + \Sigma'_{r,-r}\Sigma^{-1}_{-r,-r}\Sigma_{r,-r} \quad ,$$

where the dimension of $\boldsymbol{\mu}_{it,r}$ is $(r \times 1)$, Σ_r is dimension $(r \times r)$, and the indices *r* and *–r* refer to censored and positive elements, respectively (Huang, 2001). The fully augmented **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 θ and Σ are unchanged from those used in the pooled model. The priors are assumed independent and of the following form:

$$\pi(\theta) \sim N_K\left(\beta_0, B_0^{-1}\right) \ , \tag{5.40}$$

$$\pi(\Sigma) \sim IW_J\left(\rho_0, R_0\right) \ , \tag{5.41}$$

where $\pi(\theta)$ is a *K*-dimension multivariate normal distribution with mean β_0 and precision matrix B_0^{-1} and $\pi(\Sigma)$ is a *J*-dimension inverse Wishart distribution with degrees of freedom ρ_0 and scale R_0. The hyperparameters of the prior distributions (β_0, B_0^{-1}, ρ_0, R_0) are set to values that reflect very diffuse prior information. The values of β_0, B_0^{-1}, ρ_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:

$$p\left(\theta|\mathbf{z},\Sigma,\mathbf{u},W\right) \sim N_K\left(\beta_1, B_1^{-1}\right), \tag{5.42}$$

$$p\left(\Sigma|\mathbf{z},\theta,\mathbf{u},W\right) \sim IW_J\left(\rho_1, R_1\right). \tag{5.43}$$

The posterior distribution of θ is a *K*-dimension multivariate normal with mean $\beta_1 = \left(\sum_{i=1}^{N}\sum_{t=1}^{T_i} W_{it}'\Sigma^{-1}W_{it}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T_i} W_{it}'\Sigma^{-1}\mathbf{d}_{it}\right)$, covariance matrix $B_1^{-1} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T_i} W_{it}'\Sigma^{-1}W_{it}\right)^{-1}$, and $\mathbf{d}_{it} = \mathbf{z}_{it} - \mathbf{u}_i$. The posterior distribution of Σ is a J-dimension inverse Wishart with degrees of freedom $\rho_1 = J + N*T$ and scale $R_1 = \left(I_J J + \bar{S}N*T\right)/\left(J + N*T\right)$, where $\bar{S} = \frac{1}{N*T}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(\mathbf{d}_{it} - W_{it}\theta\right)\left(\mathbf{d}_{it} - W_{it}\theta\right)'$ and $T = \sum_{i=1}^{N} T_i$.

In addition to these adjustments to the posterior distributions for θ and Σ, 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:

$$\pi\left(\mathbf{u}\right) \sim N\left(\mu_0, M_0^{-1}\right), \tag{5.44}$$

$$\pi\left(V\right) \sim IW_J\left(\gamma_0, G_0\right), \tag{5.45}$$

where $\pi(\mathbf{u})$ is a univariate normal distribution with mean μ_0 and precision matrix M_0^{-1} and $\pi(V)$ is a *J*-dimension inverse Wishart distribution with degrees of freedom γ_0 and scale G_0. As with the prior distributions of θ and Σ, the hyperparameters are assumed to be known and are set to values that reflect very diffuse prior information. The values of μ_0, M_0^{-1}, γ_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:

$$p(\mathbf{u}|\mathbf{z},\theta,\Sigma,V,W) \sim N(\mu_1, M_1^2), \tag{5.46}$$

where $\mu_1 = \left(\left(\sum_t^{T_i}\mathbf{z}_{it} - \sum_t^{T_i} W_{it}\theta\right)\Sigma^{-1}\right)M_1^2$ is the mean of the posterior and $M_1^2 = \left(TI_j\Sigma^{-1} + V^{-1}\right)^{-1}$ is the variance. The posterior distribution of V is derived as follows:

$$p(V|\mathbf{z},\theta,\Sigma,\mathbf{u},W) \sim IW(\gamma_1, G_1), \tag{5.47}$$

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

The following outline of the steps of the Gibbs sampler for estimation of the random effects model is modified from that given above for the pooled model. The algorithm now 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 $\mathbf{y}_{it,r}^{*p}\left|\theta^{p-1},\Sigma^{p-1},\mathbf{u}_i,W_{it},\mathbf{y}_{it,-r}\right.$ for i=1,…,N from $TN_{(-\infty,0]}\left(\boldsymbol{\mu}_{it,r}^{p-1}+\Sigma_r^{p-1}\right)$, where $\boldsymbol{\mu}_{it,r}$ and $\boldsymbol{\Sigma}_r$ 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 $V^p\left|\mathbf{z}^p,\theta^{p-1},\Sigma^{p-1},\mathbf{u}^{p-1},W\right.$ from $IW\left(\gamma_1,G_1\right)$.

c. Draw $\mathbf{u}^p\left|\mathbf{z}^p,\theta^{p-1},\Sigma^{p-1},V^{p-1},W\right.$ from $N\left(\mu_1,M_1^2\right)$.

d. Draw $\Sigma^p\left|\mathbf{z}^p,\theta^{p-1},\mathbf{u}^p,W\right.$ from $IW_J\left(\rho_1,R_1\right)$.

e. Draw $\theta^p\left|\mathbf{z}^p,\Sigma^p,\mathbf{u}^p,W\right.$ from $N_K\left(\beta_1,B_1^{-1}\right)$.

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

5.3.3 Demand Model Specification

The reduced-form model estimated in this chapter is an unconditional, incomplete demand system.[23] Reduced-form is used here to indicate a model that is not integrable. That

[23] A complete demand system uses the assumption of weak separability of preferences, which is a necessary and sufficient condition for the existence of conditional demands. Conditional in this case refers to demand for a

is, the demand equations cannot be integrated back to a utility function that is consistent with consumer theory. There are several functional forms of incomplete demand systems that are based on theoretically-consistent preference orderings. Additionally, if welfare calculations are one of the goals of a study, integrability is a necessary requirement. However, the current study focuses on measuring the existence of food safety effects on consumer purchases of meat and poultry. Given that welfare calculations are not a goal of this research, a reduced-form incomplete demand system is sufficient to answer the research question of interest.

The choice of an incomplete demand system using quantities as the dependent variables was guided by previous research and technical complexity of model estimation. The development of the SUR tobit in previous research has focused on dependent variables that are estimated in levels rather than expenditures or shares (Huang, 2001). The reason for this may be recognition that a data augmentation method employed on shares would be challenging. Shares are bounded between zero and one, which has ramifications for the truncated distribution from which the augmented shares are drawn. The adding-up restriction imposed in complete demand systems adds another level of complexity to the estimation using expenditure shares. This restriction requires all the shares (latent and observed) to sum to one. This leads directly to the question of whether or not observed shares must be recalculated after drawing the observed shares. The use of an incomplete demand system avoids the issue of imposing adding up on the shares, but the effects of negative shares that

---

commodity conditional on the assumption of a first stage allocation of income across all the groups of commodities a consumer chooses to purchase. The utility maximization problem is then a second-stage optimization subject to a budget constraint of expenditures on the goods within the group. Demand for the commodities in the group of interest is, therefore, a function of the prices of the goods in the group and total expenditures on those goods. Weak separability is not assumed here as the model includes household income rather than group expenditures. Therefore, the SUR tobit demand is an unconditional demand model.

are bounded between zero and one on parameters estimates from the Gibbs sampler is an empirical question and is left for future research.

The model specifications for both the panel and random effects SUR tobit models follow closely the specifications used in the binary discrete choice models presented in section 2 of the previous chapter. The only difference between the discrete and marginal demand models is the use of a continuous dependent variable. The dependent variable for each of the three system equations is quality adjusted-per capita purchases of beef, pork, and poultry, respectively. The description of the quality adjusted quantities is given in Chapter 3.

The SUR tobit model specification includes own- and cross-effects for price and food safety, interaction terms between the media index and select demographic characteristics, and variables specific to both the household and time period. The a priori signs of these regressors do not differ from those outlined for the discrete choice models of the previous chapter.

5.4 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.[24] All the observations from the panel were used for each of the 3,000 households. This resulted in 119,280 observations that were used

[24] This is random sample 1 which was used in the multivariate conditional logit model estimation of chapter 4.

for estimation. Summary statistics are presented in table 5.1 for both the full dataset and the random sample.

Bayesian coefficients are typically the mean of the posterior samples. Drawing from the Bernstien-von Mises theorem, the posterior analysis presented in the following sections is given a classical statistical interpretation.[25] 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.

5.4.1 Posterior Samples and Convergence

The Gibbs sampler procedure was run for 11,000 iterations for the pooled SUR tobit model and 36,000 iterations for the random effects SUR tobit model.[26,27] The number of iterations that were run for each model must be large enough to allow for convergence to the

---

[25] The Bernstien-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).

[26] Both the pooled and the random effects SUR tobit models were estimated using MATLAB. Tests of the accuracy of the estimation code were conducted via a generated data experiment. Model estimation of predetermined parameter values was conducted using data from a known distribution. The Gibbs sampler was run for 1,000 iterations and convergence and mixing of the posterior distributions was accurate enough to recover the known parameter values.

[27] The number of total iterations run for each model depends on the number of skips needed to sufficiently reduce autocorrelation of the parameters and the number of iterations it takes to reach the appropriate posterior distribution (1,000 posterior realizations was the target for this study). Preliminary runs of both models indicated that 11,000 iterations are more than enough for the pooled model because only two skips are needed to reduce autocorrelation. However, a much longer chain is needed to get a sufficient number of posterior realizations for the random effects model because a larger number of skips were needed.

posterior distribution as well as sufficient mixing throughout the distribution. To ensure that the target distribution has been reached, the first 500 iterations are dropped from the posterior analysis. This is referred to as a burn-in period and ensures that inference is made on posterior realizations that have reached the posterior density of interest. Further convergence analysis of the posterior realizations of both the pooled and the random effects models are provided in Appendix A.

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^{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:

$$\mathrm{ACF}_L = \left(\frac{T}{T-L}\right)\frac{\sum_{t=1}^{T-L}(x_t-\bar{x})(x_{t+L}-\bar{x})}{\sum_{t=1}^{T}(x_t-\bar{x})^2} \quad , \tag{5.48}$$

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 each model. A plot of the ACF for all the parameters in the pooled model is shown in figure 5.1. The parameter with the highest level of autocorrelation determines the lag length that must be used to decrease the ACF of all the model parameters to 0.25 or less. For the pooled model, the lag length must be 2 to achieve this. While this is a relatively low number of skips, according to the ACF, it is sufficient to decrease autocorrelation to an acceptable level (Lynch, pp. 147, 2007). Figure 5.2 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 pooled model and keeping every other sampled value, 5,252 posterior realizations remain for inference. Similarly, the first 500 iterations are dropped from the random effects model and one in every 35 sampled values is kept for a posterior sample of 1,015 realizations. The results presented in the next section are based on these posterior sampled values for each model.

5.4.2 Pooled SUR Tobit Model Parameter Estimates

Results of the pooled SUR tobit model are presented in table 5.2. The means, standard deviations, and 95% confidence intervals are calculated using the 5,252 posterior realizations. The following discussion of the model parameters uses the means of the posterior distributions.

The own-price coefficients for each of the three commodities are statistically significantly different from zero and have the expected negative signs. The cross-price effects are also all statistically significantly different from zero and are negative. The own-effect of the media index on purchases of beef is not statistically significantly different from zero. The mean of the posterior is 0.001 and the confidence interval is evenly centered on zero. This suggests that food safety information on beef does not have a measurable average effect on the monthly quantity of beef purchased by households in the sample. The own-effect for the pork media index is -0.043. However, the 95% confidence interval contains zero, so this negative effect is not statistically significantly different from zero. The own-effect for poultry is -0.083 and is statistically significantly different from zero, suggesting that food safety information affects households' decision to purchase poultry. The cross-effects of food safety information are not statistically significantly different from zero for all the parameters except the effect of the pork media index on poultry purchases. The positive effect indicates that increases in the amount of pork food safety information have a positive and statistically significant effect on the quantity of poultry households purchase.

The interaction terms for education and food safety are statistically significantly different from zero in all three equations. The signs are negative for beef and pork, but positive for poultry. The interaction term of the pork media index and urban location of the household is statistically significantly different from zero and positive in sign. The interaction term of the poultry media index and the dummy variable for head of household age 55 and older is statistically significantly different from zero and positive in sign. Although some of the interaction effects are statistically significant, it is unclear if the sums of the average and the heterogeneous effects are also statistically significantly different from zero. The remaining parameters accounting for the heterogeneous effects of food safety are not statistically significantly different from zero.

The effect of a head of household having a college education is negative and statistically significant for all three commodities. The coefficient for households with children is negative and statistically significant, indicating that they buy less fresh meat and poultry relative to households without children. The effect of a head of household age 55 and older is positive and statistically significant for meat and poultry purchases. The coefficient for urban location is positive and statistically significant for poultry purchases, but not statistically significantly different from zero for beef and pork. Higher levels of household income increase the per capita amount of beef, pork, and poultry that is purchased. This effect declines as the level of household income increases for beef and pork, but is not statistically significant for poultry.

The remaining parameters in the model account for the variability that arises from time, geographic location, and race. The majority of the year and month dummy variables are statistically significantly different from zero. Geographic location has a significant effect for some of the commodities. Beef purchases are higher for households in the west and northeast relative to households in the south. Beef purchases of households in the central region are not statistically significantly different from households in the south. None of the geographic regions have a statistically significantly different effect for pork purchases relative to households in the south. Purchases of poultry are lower in the central region, but statistically significantly higher for households in the west and northeast relative to households in the south.

Several of the heterogeneous effects of household race are statistically significantly different from zero. For beef, Hispanic households are not statistically significantly different from white households. However, black, Asian, and other race households purchase statistically significantly less beef than their white counterparts. Pork purchases are statistically significantly lower for Hispanic households, but higher for each of the other races relative to white households. The statistically significant and positive signs of each race variable in the poultry equation indicate that all the races considered buy more poultry relative to white households.

5.4.3 Random Effects SUR Tobit Model Parameter Estimates

Results of the random effects SUR tobit model are presented in table 5.3. The means, standard deviations, and 95% confidence intervals are calculated using 1,015 posterior realizations. The following discussion of the model parameters refers to the means of the posterior distributions.

The own-price coefficients for each of the three commodities are statistically significantly different from zero and have the expected negative signs. The coefficients of the cross-price effects are statistically significantly different from zero with negative signs for all but one of the parameters. There is no statistically significant effect of the price of pork on the quantity of poultry households purchase. In general, the signs and significance of these parameters do not differ from those estimated for the pooled model.

The parameters measuring the own-effect of the media indices for beef, pork, and poultry are not statistically significantly different from zero for any of the commodities. The confidence intervals for each parameter include zero, indicating that commodity-specific food safety information does not impact households' purchase decisions for fresh meat and poultry. One of the parameters measuring the cross-effects of food safety information is statistically significantly different from zero. Increases in the amount of pork food safety information have a positive and statistically significant effect on the quantity of poultry households purchase. The remaining parameters are not statistically significantly different from zero, indicating that there is no measurable cross-commodity effect from food safety information for meat and poultry.

The food safety interaction terms for households with college educated heads of household and households with children are negative and statistically significantly different from zero in the beef and pork equations. This effect is not statistically significantly different from zero for poultry. The pork media index and head of household age 55 or older interaction term is negative and statistically significantly different from zero. as is the interaction effect between the pork media index and households with children. The interaction term of the poultry media index and the dummy variable for heads of household age 55 and older is also statistically significantly different from zero, but has a positive sign. The remaining food safety interaction parameters are not statistically significantly different from zero. As with the pooled model, the statistical significance of the interaction effects does not indicate if the sums of the average and the heterogeneous effects are also statistically significantly different from zero.

The effect of a head of household having a college education is negative and statistically significantly different from zero for beef. There is no statistically significant effect of a college education on purchases of pork and poultry. The coefficient for households with children is negative and statistically significant for each commodity, indicating that they buy less fresh meat and poultry relative to households without children. The effect of a head of household age 55 and older is positive and statistically significant for beef and pork purchases, but does not affect poultry purchases. The coefficient for urban location is positive and statistically significant for poultry purchases, but not statistically significantly different from zero for beef and pork. Higher levels of household income increase the per capita amount of beef and poultry that is purchased. This effect declines as the level of household income increases for beef, but is not significant for poultry. The effect of income on pork purchases is also increasing at a decreasing rate. However, the parameter for the income effect is not statistically significantly different from zero, while the quadratic income term is statistically significant.

The remaining parameters in the model account for the variability from time, geographic location, and race. Most of the year and month dummy variables are statistically significantly different from zero, especially in the pork and poultry equations. Geographic location has a statistically significant effect for beef and poultry, but not pork. Beef purchases are higher for households in the west and northeast relative to households in the south. Beef purchases of households in the central region are not statistically significantly different from households in the south. Purchases of poultry are lower in the central region, but statistically significantly higher for households in the west and northeast relative to households in the south.

Several of the heterogeneous effects of household race are statistically significantly different from zero for beef and pork. For beef, Hispanic households are not statistically significantly different from white households. However, black, Asian, and other race households purchase less beef than their white counterparts. The statistically significant and positive sign of the black race variable in the poultry equation indicates that these households buy more poultry relative to white households.

5.4.4 Elasticities

Household-level elasticities are given below for prices, income, and food safety for the various demographic subgroups. 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 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:

$$\boldsymbol{\theta} = \frac{\partial E\left(y_i^*\right)}{\partial W_i} , \tag{5.49}$$

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:

$$\mathbf{m} = \frac{\partial E\left(y_i\right)}{\partial W_i} \ , \tag{5.50}$$

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 ith household and jth equation of the pooled SUR tobit model are calculated as:

$$m_{ij} = \Phi\left(\frac{\left(\alpha_j + \mathbf{x}_{ij}\boldsymbol{\beta}_j + \mathbf{c}_i\boldsymbol{\gamma}_j\right)}{\left(\Sigma_{jj}\right)^{\frac{1}{2}}}\right)\theta_j \ , \tag{5.51}$$

where Σ_{jj} is the jth diagonal element of the covariance matrix, $\theta_j = \left[\alpha_j \ \boldsymbol{\beta}_j \ \boldsymbol{\gamma}_j\right]$ are the parameter estimates for the jth equation, and $\Phi\left(\cdot\right)$ is the standard normal cdf. The marginal effects are modified slightly for the random effects model to reflect the additional variance of the household-specific error component. The marginal effects for the ith household and the jth equation of the random effects SUR tobit model are calculated as:

$$m_{ij} = \Phi\left(\frac{\left(\alpha_j + \mathbf{x}_{ij}\boldsymbol{\beta}_j + \mathbf{c}_i\boldsymbol{\gamma}_j\right)}{\left(\Sigma_{jj} + V_{jj}\right)^{\frac{1}{2}}}\right)\theta_j \ , \tag{5.52}$$

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

The own-price elasticity of the jth commodity is calculated as follows:

$$E_j^{price} = m_j^{price} * \frac{\overline{p}_j}{\overline{y}_j} \quad , \tag{5.53}$$

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

$$E_{jl}^{price} = m_{jl}^{price} * \frac{\overline{p}_l}{\overline{y}_j} \text{ for } j \neq l \,, \tag{5.54}$$

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

$$E_j^{inc} = \left(m_j^{inc} + 2 * m_j^{inc^2} * \overline{inc} \right) * \frac{\overline{inc}}{\overline{y}_j} \quad , \tag{5.55}$$

where m_j^{inc} is the income marginal effect for the jth commodity, $m_j^{inc^2}$ is the income squared marginal effect for the jth 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:

$$E_j^{MI*Ed} = \left(m_j^{MI} + m_j^{MI*Ed}\right) * \frac{\overline{MI}_j}{\bar{y}_j^{Ed}} \ , \qquad (5.56)$$

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, income, and food safety elasticities for the pooled model are presented in table 5.5. 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 poultry, indicating that consumer response to a price change is relatively more elastic for these products as compared to pork. The elasticity for beef suggests that a 10% increase in the price of beef would cause a 14.7% decline in per capita beef purchases. The effect from a 10% increase in the price of pork is estimated to be a 8.6% decline in purchases. The price effect for poultry is the most elastic with an estimated decrease of 18.5% from a 10% increase in price. The cross-price elasticities are also statistically different from zero and have negative signs. These cross-price effects are inelastic compared to the own-price effects 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 2.8%. The effects for pork and poultry are increases in per capita purchases of 1.7% and 2.4%, respectively.

Seven of the twelve food safety elasticities are statistically significantly different from zero using a 95% confidence interval. The interaction effect of the food safety media index and the dummy variable for households with college educated heads is statistically significantly different from zero for beef, pork, and poultry. For households with a college educated head, a 10% increase in the media index for beef will decrease per capita purchases by 0.2%. The effect for these households from a 10% increase in the pork media index is a 0.3% decrease in pork purchases and a 10% increase in the media index for poultry will cause a 0.4% decrease in per capita purchases of poultry. The effect of food safety information on households with children present is statistically significantly different from zero for pork and poultry, but not beef. A 10% increase in the pork media index will decrease per capita purchases of pork by 0.3%, while an increase in the poultry media index is expected to decrease poultry purchases by 0.3%. The effect of both the pork and poultry food safety media indices on purchases is also statistically significant for households located in urban areas. The estimated decrease for these households is 0.5% for a 10% increase in the poultry media index. However, the effect of a 10% increase in pork media index is an increase of 0.12% for pork purchases. The remaining food safety elasticities are not statistically significantly different from zero.

The elasticity results suggest that there are some groups of consumers that are influenced by the food safety information available in the media, especially households with college educated heads and those with children. Households with heads age 55 and older do not appear to react to food safety information any differently than the average household in the data sample. Although some of the food safety effects are statistically significant for some households, the relative magnitude of these elasticities as compared to the price and income elasticities is small. This suggests that food safety information does not necessarily have an economically significant effect on the amount of meat and poultry household purchase.

The price and food safety elasticities are presented in table 5.6 for the random effects model. As in the pooled 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. As with the pooled model, 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 of 0.4% for these households. However, an increase in the beef and pork media indices is estimated to cause a 0.12% increase in the amount of beef and pork urban household purchase. All the remaining food safety elasticities are not significantly different from zero. As in the pooled model, 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 both the pooled and random effects models for beef and pork fall within these ranges.[28] 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 comparable to the food safety elasticities found in this study for each of the four demographic groups of households.

5.5 Conclusion

The results of the pooled model suggest that an increase in the amount of food safety information regarding poultry negatively impacts the amount of poultry purchased by households with college educated heads, household with children present, and urban household. Beef purchases are also negatively impacted by beef food safety information for households with college educated heads of household. The pork media index negatively impacts households with college educated heads and children present, but positively impacts pork purchases by urban households. There is no statistically significant effect from beef food safety information for urban households. Households with heads age 55 and older are

[28] 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.

also not expected to react to food safety information differently than the average household for any of the media indices.

Although several statistically significant heterogeneous food safety effects were found using the pooled SUR tobit model, the elasticities calculated from the results of the random effects SUR tobit model indicate that food safety 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. These effects were positive for beef and pork food safety information and negative for poultry information. 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.

Schlenker and Villas-Boas (2006) found statistically significant effects at the grocery store level, but not at the household level for purchases of meat. 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 lack of statistical significance of most of the food safety variables in the random effects model suggests that fitting the panel data more precisely using a component error structure has implications for the conclusions drawn about these second-order effects. The results of the pooled model indicate that, although small in magnitude, the majority of households do respond to food safety information. However, once unobserved heterogeneity at the household level is accounted for by a component error, the statistical significance of most of these effects no longer remains and the magnitude of those that do are even smaller. The difference in conclusions that can be drawn from the results of the two SUR tobit models provides some justification for explicitly modeling a component error structure.

Table 5.1 Summary Statistics of Demand Model Variables

	Full Sample				Random Sample			
	Average	**Minimum**	**Maximum**	**Std. Dev.**	**Average**	**Minimum**	**Maximum**	**Std. Dev.**
Beef Price	3.209	0.577	12.638	0.562	3.196	1.227	12.638	0.551
Pork Price	2.534	0.627	12.219	0.509	2.527	0.644	11.453	0.513
Poultry Price	1.924	0.700	8.195	0.248	1.918	0.880	7.082	0.248
Beef MI	7.633	0.786	77.645	6.428	7.650	0.786	77.645	6.446
Pork MI	2.547	0.000	16.567	1.988	2.558	0.000	16.567	2.010
Poultry MI	11.378	2.000	38.310	6.054	11.336	2.000	38.310	6.021
Ed	0.393	0	1	0.488	0.376	0	1	0.484
Age	0.372	0	1	0.483	0.376	0	1	0.484
Urban	0.875	0	1	0.330	0.873	0	1	0.333
Child	0.296	0	1	0.456	0.288	0	1	0.453
Income	5.383	0.250	12.500	3.151	5.281	0.250	12.500	3.137
Income^2	38.910	0.062	156.250	43.477	37.729	0.062	156.250	43.064
Y1	0.120	0	1	0.325	0.120	0	1	0.325
Y2	0.112	0	1	0.316	0.114	0	1	0.318
Y3	0.118	0	1	0.322	0.118	0	1	0.323
Y4	0.127	0	1	0.333	0.130	0	1	0.337
Y5	0.133	0	1	0.340	0.131	0	1	0.338
Y6	0.136	0	1	0.342	0.134	0	1	0.341
Y7	0.129	0	1	0.336	0.130	0	1	0.336
Y8	0.125	0	1	0.330	0.122	0	1	0.328
M1	0.083	0	1	0.276	0.083	0	1	0.276
M2	0.083	0	1	0.276	0.083	0	1	0.276
M3	0.083	0	1	0.276	0.083	0	1	0.276
M4	0.083	0	1	0.276	0.083	0	1	0.276
M5	0.083	0	1	0.276	0.083	0	1	0.276
M6	0.083	0	1	0.276	0.083	0	1	0.276
M7	0.083	0	1	0.276	0.083	0	1	0.276
M8	0.083	0	1	0.276	0.083	0	1	0.276
M9	0.083	0	1	0.276	0.083	0	1	0.276
M10	0.083	0	1	0.276	0.083	0	1	0.276
M11	0.083	0	1	0.276	0.083	0	1	0.276
M12	0.083	0	1	0.276	0.083	0	1	0.276
South	0.366	0	1	0.482	0.362	0	1	0.481
Central	0.204	0	1	0.403	0.216	0	1	0.412
West	0.217	0	1	0.412	0.216	0	1	0.412
Northeast	0.213	0	1	0.410	0.205	0	1	0.404
Caucasian	0.766	0	1	0.423	0.758	0	1	0.429
Hispanic	0.076	0	1	0.264	0.075	0	1	0.264
Black	0.121	0	1	0.326	0.123	0	1	0.328
Asian	0.022	0	1	0.146	0.026	0	1	0.159
Other	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 5.2 Bayesian Estimated Coefficients of the Pooled SUR Tobit Model

	Beef Model				Pork Model				Poultry Model			
	Coefficent	Standard Deviation	95% Confidence Interval		Coefficent	Standard Deviation	95% Confidence Interval		Coefficent	Standard Deviation	95% Confidence Interval	
Beef Price	-8.498	0.117	-8.733	-8.263	-1.429	0.097	-1.629	-1.235	-0.735	0.098	-0.930	-0.547
Pork Price	-0.946	0.125	-1.193	-0.701	-6.391	0.090	-6.570	-6.212	-0.206	0.101	-0.405	-0.011
Poultry Price	-2.811	0.237	-3.270	-2.353	-2.571	0.189	-2.941	-2.203	-15.933	0.179	-16.286	-15.585
Beef MI	0.001	0.023	-0.045	0.046	-0.002	0.007	-0.017	0.012	0.006	0.007	-0.009	0.021
Pork MI	0.019	0.031	-0.041	0.079	-0.043	0.059	-0.158	0.074	0.063	0.024	0.015	0.109
Poultry MI	-0.004	0.011	-0.026	0.019	-0.017	0.009	-0.035	0.000	-0.083	0.023	-0.128	-0.038
Ed*MI_{beef}	-0.055	0.013	-0.082	-0.029	--	--	--	--	--	--	--	--
Age*MI_{beef}	-0.017	0.014	-0.045	0.011	--	--	--	--	--	--	--	--
Child*MI_{beef}	0.002	0.016	-0.028	0.034	--	--	--	--	--	--	--	--
Urban*MI_{beef}	0.019	0.020	-0.020	0.058	--	--	--	--	--	--	--	--
Ed*MI_{pork}	--	--	--	--	-0.130	0.035	-0.197	-0.060	--	--	--	--
Age*MI_{pork}	--	--	--	--	0.005	0.038	-0.070	0.080	--	--	--	--
Child*MI_{pork}	--	--	--	--	-0.077	0.039	-0.154	0.000	--	--	--	--
Urban*MI_{pork}	--	--	--	--	0.126	0.052	0.020	0.225	--	--	--	--
Ed*$MI_{poultry}$	--	--	--	--	--	--	--	--	0.025	0.012	0.002	0.048
Age*$MI_{poultry}$	--	--	--	--	--	--	--	--	0.060	0.013	0.035	0.086
Child*$MI_{poultry}$	--	--	--	--	--	--	--	--	0.002	0.014	-0.026	0.029
Urban*$MI_{poultry}$	--	--	--	--	--	--	--	--	0.015	0.020	-0.023	0.054
Ed	-1.681	0.137	-1.956	-1.410	-1.262	0.116	-1.495	-1.036	-0.791	0.152	-1.085	-0.492
Age	1.388	0.145	1.111	1.669	1.809	0.124	1.566	2.049	0.024	0.165	-0.298	0.348
Child	-3.742	0.160	-4.058	-3.431	-2.126	0.132	-2.384	-1.865	-2.583	0.176	-2.933	-2.239
Urban	-0.017	0.196	-0.406	0.368	-0.297	0.167	-0.622	0.041	1.113	0.233	0.646	1.564
Income	1.394	0.052	1.288	1.494	0.827	0.042	0.745	0.911	0.732	0.042	0.649	0.815
$Income^2$	-0.045	0.004	-0.052	-0.038	-0.023	0.003	-0.029	-0.018	-0.002	0.003	-0.008	0.003

Note: The estimated coefficients are means calculated from 5,252 posterior realizations. The 95% confidence intervals are calculated using the upper and lower 2.5% tails of the posterior distribution.

Table 5.2 Bayesian Estimated Coefficients of the Pooled SUR Tobit Model, cont.

	Beef Model				Pork Model				Poultry Model			
	Coefficent	Standard Deviation	95% Confidence Interval		Coefficent	Standard Deviation	95% Confidence Interval		Coefficent	Standard Deviation	95% Confidence Interval	
Y1	-0.018	0.216	-0.442	0.407	7.111	0.171	6.775	7.439	4.300	0.177	3.945	4.647
Y2	-2.499	0.180	-2.848	-2.143	0.501	0.144	0.215	0.784	-0.248	0.150	-0.544	0.045
Y3	-0.660	0.166	-0.996	-0.329	0.520	0.141	0.242	0.789	0.805	0.140	0.517	1.074
Y4	-0.334	0.177	-0.685	0.012	0.984	0.142	0.710	1.266	0.104	0.146	-0.177	0.398
Y5	-0.653	0.164	-0.973	-0.332	-0.269	0.127	-0.519	-0.020	-0.653	0.133	-0.911	-0.388
Y7	1.968	0.166	1.637	2.285	1.220	0.134	0.962	1.492	1.468	0.138	1.208	1.743
Y8	0.734	0.167	0.405	1.058	1.175	0.140	0.905	1.446	2.159	0.140	1.881	2.427
M1	-0.322	0.220	-0.762	0.116	-2.492	0.170	-2.825	-2.150	0.999	0.177	0.648	1.350
M2	-0.495	0.206	-0.904	-0.098	-2.682	0.162	-2.993	-2.366	0.920	0.165	0.603	1.248
M3	0.174	0.206	-0.211	0.574	-2.020	0.161	-2.341	-1.705	1.018	0.167	0.696	1.355
M4	-0.297	0.208	-0.706	0.102	-1.413	0.158	-1.723	-1.096	0.702	0.165	0.381	1.035
M5	1.398	0.208	0.995	1.815	-2.305	0.161	-2.615	-1.991	1.501	0.165	1.172	1.819
M6	0.536	0.208	0.133	0.937	-2.695	0.163	-3.016	-2.374	1.013	0.169	0.679	1.340
M7	0.539	0.212	0.121	0.955	-2.473	0.165	-2.800	-2.149	1.115	0.172	0.776	1.451
M8	0.694	0.210	0.286	1.101	-2.421	0.163	-2.741	-2.101	1.616	0.166	1.283	1.942
M9	0.184	0.205	-0.222	0.577	-2.466	0.164	-2.792	-2.144	1.182	0.168	0.850	1.513
M10	0.121	0.203	-0.283	0.517	-2.505	0.161	-2.818	-2.199	0.786	0.165	0.471	1.114
M11	-1.633	0.215	-2.060	-1.214	-2.262	0.162	-2.587	-1.947	0.962	0.169	0.631	1.291
Central	0.236	0.136	-0.032	0.508	-0.097	0.111	-0.322	0.121	-1.373	0.113	-1.591	-1.156
West	2.039	0.150	1.748	2.344	-0.262	0.127	-0.504	-0.004	2.041	0.125	1.797	2.286
Northeast	0.718	0.127	0.473	0.969	-0.016	0.104	-0.219	0.186	0.801	0.105	0.596	1.014
Hispanic	0.177	0.157	-0.132	0.483	-0.428	0.136	-0.699	-0.164	0.560	0.131	0.301	0.811
Black	-2.751	0.133	-3.012	-2.496	1.394	0.102	1.192	1.593	2.590	0.103	2.388	2.789
Asian	-3.396	0.283	-3.968	-2.853	0.953	0.219	0.517	1.389	0.952	0.214	0.528	1.365
Other	-2.140	0.315	-2.756	-1.529	1.249	0.248	0.786	1.752	0.491	0.249	0.003	0.982
Constant	28.373	0.584	27.216	29.549	18.513	0.459	17.614	19.423	23.350	0.478	22.393	24.283
Sigma$_\varepsilon$	14.882	0.041	14.800	14.964	10.310	0.038	10.234	10.385	11.376	0.036	11.305	11.446

Note: The estimated coefficients are means calculated from 5,252 posterior realizations. The 95% confidence intervals are calculated using the upper and lower 2.5% tails of the posterior distribution.

Table 5.3 Bayesian Estimated Coefficients of the Random Effects SUR Tobit Model

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

Table 5.3 Bayesian Estimated Coefficients of the Random Effects SUR Tobit Model, cont.

	Beef Model				Pork Model				Poultry Model			
	Coefficent	Standard Deviation	95% Confidence Interval		Coefficent	Standard Deviation	95% Confidence Interval		Coefficent	Standard Deviation	95% Confidence Interval	
Y1	0.265	0.223	-0.199	0.675	6.617	0.181	6.252	6.969	4.383	0.176	4.030	4.727
Y2	-1.453	0.179	-1.791	-1.087	1.170	0.150	0.870	1.468	0.442	0.156	0.134	0.748
Y3	-0.081	0.170	-0.404	0.260	0.868	0.142	0.588	1.133	1.207	0.137	0.930	1.480
Y4	-0.331	0.166	-0.661	0.003	0.897	0.137	0.640	1.188	0.349	0.130	0.084	0.597
Y5	-0.713	0.146	-1.010	-0.429	-0.200	0.122	-0.435	0.045	-0.452	0.124	-0.701	-0.225
Y7	1.109	0.159	0.808	1.414	0.487	0.129	0.217	0.731	0.724	0.123	0.461	0.952
Y8	0.043	0.173	-0.299	0.392	0.513	0.140	0.222	0.785	1.352	0.131	1.090	1.600
M1	-0.148	0.196	-0.533	0.240	-2.269	0.167	-2.587	-1.938	0.967	0.158	0.670	1.277
M2	-0.407	0.190	-0.778	-0.048	-2.635	0.155	-2.948	-2.349	0.823	0.148	0.528	1.118
M3	0.328	0.187	-0.043	0.679	-1.812	0.147	-2.105	-1.527	0.945	0.150	0.661	1.243
M4	-0.268	0.186	-0.633	0.104	-1.326	0.147	-1.617	-1.035	0.589	0.148	0.296	0.875
M5	1.335	0.191	0.975	1.696	-2.466	0.148	-2.743	-2.174	1.298	0.153	0.998	1.591
M6	0.479	0.192	0.096	0.866	-2.863	0.154	-3.179	-2.560	0.841	0.151	0.530	1.130
M7	0.457	0.191	0.080	0.830	-2.697	0.149	-2.996	-2.407	0.909	0.158	0.594	1.211
M8	0.525	0.184	0.173	0.888	-2.703	0.151	-2.982	-2.401	1.291	0.151	1.009	1.594
M9	0.141	0.188	-0.234	0.502	-2.603	0.148	-2.893	-2.326	0.988	0.149	0.698	1.261
M10	0.143	0.187	-0.243	0.496	-2.511	0.148	-2.798	-2.230	0.687	0.146	0.402	0.985
M11	-1.335	0.190	-1.685	-0.965	-1.974	0.152	-2.266	-1.687	1.529	0.145	1.261	1.809
Central	0.084	0.440	-0.776	0.966	0.184	0.350	-0.481	0.880	-1.070	0.336	-1.739	-0.444
West	1.622	0.435	0.804	2.539	-0.535	0.329	-1.158	0.102	1.992	0.329	1.394	2.678
Northeast	1.391	0.425	0.556	2.226	0.296	0.303	-0.330	0.882	1.284	0.293	0.704	1.879
Hispanic	0.703	0.422	-0.110	1.557	0.086	0.335	-0.548	0.733	0.596	0.311	0.033	1.216
Black	-2.306	0.455	-3.187	-1.344	0.502	0.347	-0.209	1.165	1.725	0.329	1.086	2.381
Asian	-2.180	0.667	-3.476	-0.857	0.326	0.520	-0.783	1.297	0.428	0.496	-0.579	1.392
Other	-2.502	0.489	-3.478	-1.558	0.749	0.400	-0.025	1.538	0.021	0.380	-0.718	0.786
Constant	22.693	0.744	21.164	24.127	10.493	0.630	9.274	11.694	17.886	0.657	16.633	19.156
Sigma$_\varepsilon$	12.363	0.034	12.296	12.433	8.695	0.032	8.632	8.756	9.480	0.030	9.421	9.538
Sigma$_u$	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 5.4 Price and Food Safety Elasticities of Pooled SUR Tobit Models

		Elasticity	Standard Deviation	95% Confidence Interval	
Own-Price					
	Beef	-1.470	0.020	-1.509	-1.430
	Pork	-0.856	0.012	-0.880	-0.831
	Poultry	-1.851	0.021	-1.892	-1.809
Cross-Price					
Beef	**Pork**	-0.134	0.018	-0.168	-0.099
	Poultry	-0.290	0.025	-0.340	-0.242
Pork	**Beef**	-0.236	0.016	-0.267	-0.205
	Poultry	-0.253	0.018	-0.289	-0.217
Poultry	**Beef**	-0.143	0.019	-0.181	-0.106
	Pork	-0.033	0.016	-0.064	-0.003
Income					
	Beef	0.275	0.006	0.263	0.287
	Pork	0.167	0.005	0.158	0.176
	Poultry	0.239	0.006	0.228	0.250
Food Safety					
College Education	**Beef**	-0.025	0.011	-0.045	-0.004
	Pork	-0.025	0.009	-0.042	-0.008
	Poultry	-0.043	0.017	-0.077	-0.010
Age 55 & Older	**Beef**	-0.005	0.008	-0.021	0.010
	Pork	-0.005	0.007	-0.018	0.008
	Poultry	-0.014	0.013	-0.040	0.012
Children Present	**Beef**	0.002	0.015	-0.027	0.032
	Pork	-0.028	0.014	-0.055	-0.002
	Poultry	-0.093	0.027	-0.146	-0.040
Urban Residence	**Beef**	0.009	0.006	-0.003	0.021
	Pork	0.012	0.005	0.002	0.022
	Poultry	-0.050	0.010	-0.069	-0.031

Note: The own- and cross-price elasticities are means calculated from 5,252 posterior realizations. The 95% confidence intervals are calculated using the upper and lower 2.5% tails of the posterior distribution.

Table 5.5 Price and Food Safety Elasticities of Random Effects SUR Tobit Models

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

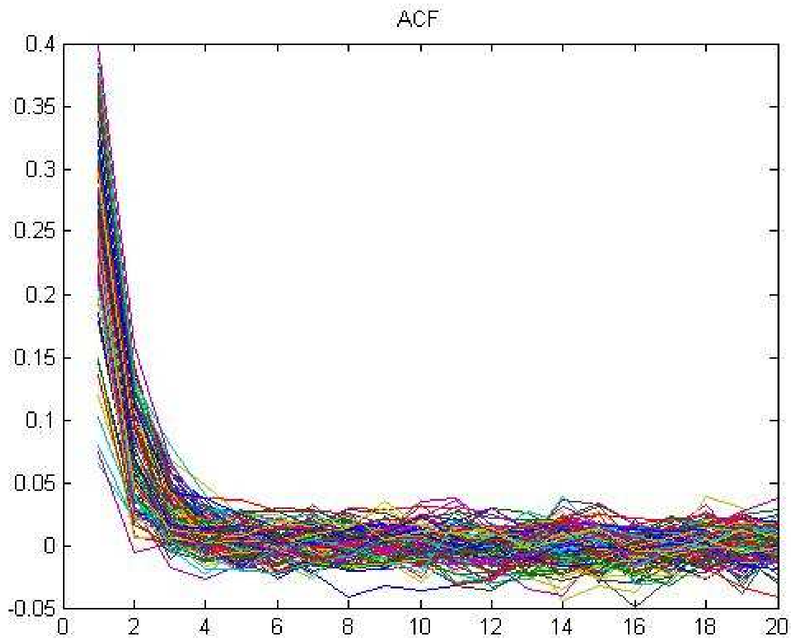

Figure 5.1 ACF at Different Lag Lengths of all the Parameters in the Pooled SUR Tobit Model

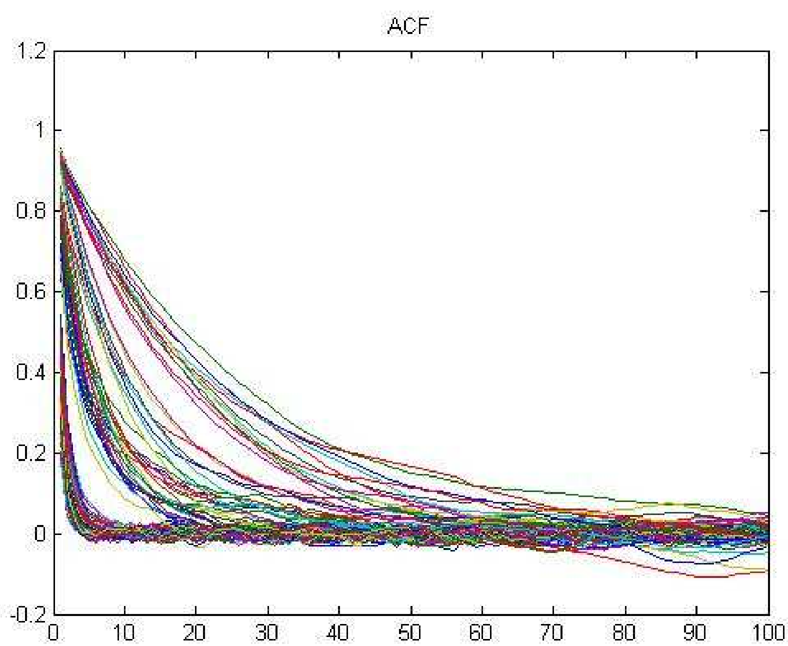

Figure 5.2 ACF at Different Lag Lengths of all the Parameters in the Random Effects SUR Tobit Model

6 Chapter

Conclusions and Future Research

6.1 Conclusion

The research presented in this study had two primary objectives. The first objective was to determine if the effects of food safety information varies across heterogeneous consumers. Results from both the discrete choice models and the pooled SUR tobit demand model indicate that food safety information may impact consumer decisions to purchase meat and poultry. However, the random effects SUR tobit model results indicate that, once the correlation between purchases by the same household is accounted for using a component error structure, very little evidence of food safety impacts remain.

The second objective of this study was to develop an estimation strategy that would account for several aspects of household-level panel data in the same model. The model presented in chapter 5 simultaneously incorporates censoring, correlation between observations from the same household, and correlation between the error terms of the demand equations.

Microeconomic data is a rich source of information on consumer choice, but it poses many econometric challenges to the researcher. The censored observations are as interesting as the purchase observations, especially when considering the effects of food safety information on whether or not consumers choose to keep making purchases of meat and poultry. This study contributes to the existing literature by analyzing the effects of food safety information on heterogeneous households using an estimation strategy that accounts econometrically for censored panel data in a demand system. The use of household level data provides the opportunity to examine how the heterogeneity of consumers impacts their response to publically available food safety information, which has not been previously analyzed in a demand system framework. The SUR tobit models used in this study have several practical advantages that will be useful for future analysis using household level panel data. The models are relatively straightforward to implement, are not limited in the number of censored equations that can be included in the demand system, and can be modified to include a component error term.[29] Therefore, these models provide a good base from which further advances in the estimation of demand systems using microeconomic data can be made.

6.1.1 Discrete Choice Models of Meat and Poultry Purchases

The objective of the research presented in chapter 4 was to investigate if the quantity of food safety information available to consumers impacts their purchase decisions for fresh meat and poultry using a discrete choice estimation framework. The food safety information

[29] Although the Gibbs sampler methodology is straightforward to implement, some models will take longer to estimate than others. In this study, the random effects model took approximately three times longer to estimate than the pooled model. The amount of time required to generate a sufficient number of posterior realizations to perform a reliable analysis is a function of several factors including: functional form of the model, size of the dataset, coding efficiency, and level of autocorrelation among the model parameters, to name a few.

index used in the models represents the level of food safety information available to the public from regional newspapers. The media index was interacted with various demographic characteristics within in the model to determine if the effect of food safety information varies across different groups of households.

Both binary and multinomial logit models were estimated to investigate the effects of the different types of food safety information on purchase decisions. Results from estimation of the commodity-specific binary choice models suggest that responses to food safety information do vary across households for beef and pork, while poultry purchase probabilities are not affected. There were a few unexpected results from the binary choice models that prompted further investigation of the meat and poultry purchase decisions in a more complex multinomial choice model that accounts for the interactions between purchase alternatives. The specification of the 8-choice logit model was unique in the grouping of explanatory variables to isolate effects of the price, food safety information, and household characteristics into commodity-specific effects. Interaction terms were again included to investigate any effects from food safety information that are specific to certain groups of households and may differ from the average effect across the entire population. The results of the 8-choice model suggest that the households most likely to stop purchasing beef in a given month, when the amount of food safety information increases, are those with a college educated head of household. This avoidance behavior is also present for households in urban areas, with respect to poultry purchases. Other households do not appear to have a measurable response to food safety information with regard to discrete purchase decisions of beef, pork, and poultry.

6.1.2 Demand Models of Meat and Poultry Consumption

The results of the pooled SUR tobit model suggest that an increase in the amount of food safety information on poultry impacts the quantity of poultry purchased. For households with a college educated head of household, food safety information on all three commodities impacts purchase decisions. Food safety information on pork and poultry affects those households with children present, while households located in urban areas are only impacted by poultry food safety information. There is no statistically significant effect from beef or pork food safety information for urban households and households with heads age 55 and older are also not expected to react to food safety information for any of the three commodities.

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. Households located in urban areas have a statistically significant response which is negative for poultry purchases, but the response is positive for beef and pork. 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.

The notable differences in results between the pooled and random effects SUR tobit models suggests that the explicitly accounting for the unobserved, household-specific component of the error structure is important. In fact, very different conclusions may be drawn regarding the effects of food safety information on household purchases of meat and poultry depending on which model is estimated. These results provide some justification for employing a more computationally intensive estimation technique in order to avoid inaccurately measuring food safety information effects.

6.2 Future Research

The research presented in this dissertation provides a strong base from which future work may emanate. Areas of future research include refinement of model and parameter specification, consideration of alternative estimation techniques, and estimation of a structural demand model. The following discussion provides an outline of those directions for future research.

6.2.1 Model and Variable Specification

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 can be captured using state dependence variables. State dependence can capture both inventory and purchase habit effects. An increase in a consumer's propensity to repurchase is referred to as habit formation, while a decrease in their probability to repurchase may be considered variety-seeking (Moeltner and Englin, 2004). Recent research by Zhen and Wohlgenant (2006) indicates that habit formation may alter the impact of food safety information on meat and poultry demand. By explaining the variability due to state dependence, second-order effects from food safety information may be more accurately identified.

There are a variety of ways to specify a media index of food safety information. 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.

Another issue with specification of the media index is which articles to include. 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.

The specification of the quality-adjusted price index may also be revisited in future research. The current price indices are specific to the four geographic regions given in the Nielsen data. However, it is reasonable to assume that geographic differences due to transportation costs, for example, would vary between urban and rural areas as well as across different regions of the United States. Therefore, the indices could further account for price variability if they were expanded to include an urban and non-urban version of each of the four regional indices.

Another area for future research is to change the level of aggregation on household purchases. Results from previous studies found different effects of food safety at the grocery store versus household level (Schlenker and Villas-Boas, 2006). The lack of evidence of food safety effects found in this study may be further supported or possibly refuted if alternative aggregation levels are used to perform the analysis. Either outcome would be an important contribution to the current food safety literature.

6.2.2 Alternative Estimation Techniques

As mentioned in chapter 4, employing a logit model of discrete choice imposes the proportional substitution patterns due to the property of *independence from irrelevant* *alternatives* (IIA). Future research that focuses on the cross-elasticities or policy analysis using welfare calculations will require the restriction of proportional substitution to be lifted. This can be accomplished by estimating a random parameters (mixed) logit model. Estimation of a mixed logit model may be computationally burdensome given the large sample of data used in this study. Therefore, one strategy that could be employed would be stratification of the data into various demographic subgroups. The subgroups would reflect the current model specification and would be determined by education and age of the head of household, presence of children in the household, and location of the household in an urban area. Estimation of the mixed logit model would be conducted using the subgroups and the impacts of food safety information on purchase decisions for these specific households could be measured.

6.2.3 Structural Demand Models

Another area for future research is specification of a structural tobit demand model rather than the reduced form specification used in this study. A structural demand model that meets certain integrability requirements would have underlying preferences that are consistent with consumer theory. However, estimation of a structural demand model is challenging when using a data augmentation estimation technique because cross-equation restrictions must be imposed. For example, a complete demand system that assumes weak separability of preferences is estimated using group expenditure shares as the dependent variable. Shares are defined using total quantity purchased, price, and total expenditures on all the items included in the demand system. The adding-up restriction is imposed on the shares and requires them to sum to one. If one or more of the items in the demand system is not purchased, then the observed share is equal to zero and the remaining items have shares that sum to one. Data augmentation, however, fills in the censored observations from a normal distribution that is truncated at zero, thereby resulting in negative expenditure shares. Imposing the adding-up restriction is further complicated by these negative shares as well as the shares of the non-censored items that would have to be greater than one to satisfy the adding-up restriction.

Future research will attempt to address this issue by using an incomplete demand system, which does not require an adding up restriction. There are several functional forms that can be used to estimate an incomplete demand system and it is important to consider the tradeoffs of the various choices. The linear-quadratic incomplete demand system (LQ-IDS) functional form is considerably less restrictive than other incomplete demand system functional forms. For example, functional forms that are linear or semi-log in quantity, price, and income do not allow for flexibility of the income and Marshallian cross-price effects when Slutsky symmetry is imposed (von Haefen, 2002). The individual income coefficients of the LQ-IDS, however, may be positive, negative, or zero and the matrix of price effects is not necessarily symmetric (LaFrance, 1990).

Another area of future interest that results from this study is the specification of a model of household demand that accounts more fully for the economic reasons censored observations are observed in household data. If household budgets are binding, then non-purchase could be due to prices that are high enough to keep people from buying meat and poultry. Decisions of how much to purchase in a given period may also be affected by the inventory of meat and poultry that households are currently holding. Future research will focus on corner solution models, such as the model put forth by Lee and Pitt (1986), that also account for household inventory levels of meat and poultry.

Bibliography

Brown, J. "Effect of a Health Hazard "Scare" on Consumer Demand." *American Journal of Agricultural Economics*. 51(1969): 676-678.

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.

Cameron, T.A. "The Impact of Grouping Coarseness in Alternative Grouped-Data Regression Models." *Journal of Econometrics.* 35(1987): 37-57.

Capps, O. Jr. and J. Park. "Impacts of Advertising, Attitudes, Lifestyles, and Health on the Demand for U.S. Pork: A Micro-Level Analysis." *Journal of Agricultural and Applied Economics.* 34(2002): 1-15.

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.

Crump, J.A., P.M. Griffin, and F.J. Angulo. "Bacterial Contamination of Animal Feed and Its Relations to Human Foodborne Illness." *Food Safety.* 35(2002).

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.

Deaton, A. "Quality, Quantity, and Spatial Variation of Price." *American Economic Review*. 78(1989): 418-430.

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.

Dong, D., J. Shonkwiler, and O. Capps, Jr. "Estimation of Demand Functions Using Cross-Sectional Household Data: The Problem Revisited." *American Journal of Agricultural Economics.* 80(1998): 466-473.

Erdem, T., M.P. Keane, and B. Sun. "Missing Price and Coupon Availability Data in Scanner Panels: Correcting for the Self-Selection Bias in Choice Model Parameters." *Journal of Econometrics.* 89(1999): 177-196.

Gao, X.M. and T. Spreen. "A Microeconomic Analysis of the U.S. Meat Demand." *Canadian Journal of Agricultural Economics.* 42(1994): 397-412.

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.

Heckman, J.J. "Sample Selection Bias as a Specification Error." *Econometrica.* 47(1979): 153-61.

Heien, D. and C.R. Wessells. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *Journal of Business and Economic Statistics.* 8(1990): 365-371.

Huang, H.-C. "Bayesian Analysis of the SUR Tobit Model." *Applied Economics Letters.* 8(2001): 617-622.

Huang, C.J., F.A. Sloan, and K.W. Adamache. "Estimation of Seemingly Unrelated Tobit Regressions via the EM Algorithm." *Journal of Business and Economic Statistics.* 5(1987): 425-430.

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.

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

Krinsky, I. and A.L. Robb. “Three Methods for Calculating the Statistical Properties of Elasticities: A Comparison.” *Empirical Economics.* 16(1991): 199-209.

LaFrance, J.T. “Incomplete Demand Systems and Semilogarithmic Demand Models.” *Australian Journal of Agricultural Economics.* 34(1990): 118-131.

Lee, L.F. and M.M. Pitt. “Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach.” *Econometrica.* 54(1986): 1237-1242.

Lusk, J. and T. Schroeder. “Effects of Meat Recalls on Futures Market Prices.” Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL, April 2000.

Lynch, S.M. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer, 2007.

Marsh, T.L, T.C. Schroeder, and J. Mintert. “Impacts of Meat Product Recalls on Consumer Demand in the USA.” *Applied Economics.* 36(2004): 897-909.

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

Meyerhoefer, C.D., C.K. Ranney, and D.E. Sahn. “Consistent Estimation of Censored Demand Systems using Panel Data.” *American Journal of Agricultural Economics*. 87(2005): 660-672.

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.

Nelson, J. “Quality Variation and Quantity Aggregation in Consumer Demand for Food.” *American Journal of Agricultural Economics*. 73(1991): 1204-1212.

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/

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.

Shonkwiler, J.S. and S.T. Yen. "Two-Step Estimation of a Censored System of Equations." *American Journal of Agricultural Economics.* 81(1999): 972-982.

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.

Stewart, M.B. "On Leas Squares Estimation when the Dependent Variable is Grouped." *Review of Economic Studies.* (1983): 737-753.

Thomsen, M. and A. McKenzie. "Market Incentives for Safe Foods: An Examination of Shareholder Losses from Meat and Poultry Recalls." *American Journal of Agricultural Economics.* 83(2001): 526-538.

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.

Von Haefen, R.H. "A Complete Characterization of the Linear, Log-Linear, and Semi-Log Incomplete Demand System Models." *Journal of Agricultural and Resource Economics*. 27(2002): 281-319.

Wales, T.J. and A.D. Woodland. "Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints." *Journal of Econometrics.* 21(1983): 263-85.

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

Zhen, C. and M.K. Wohlgenant. “Meat Demand Under Rational Habit Persistence.” *Canadian Journal of Agricultural Economics.* 54(2006): 477-495.

Appendices

Appendix A

Bayesian Estimation Issues

A.1 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 both the pooled and random effects SUR tobit models are shown in figures A.1 and A.2. 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 figures A.3 are the media index and price parameters of the pooled SUR tobit model. The histograms include only the 5,252 posterior realizations that are kept after omitting every other iteration to decrease autocorrelation. Each of the parameter histograms is very close to a normal density. This suggests that both convergence to the posterior distribution and mixing throughout this distribution have been sufficiently attained with a chain of 11,000 iterations.

The histograms for the media index and price parameters of the random effects SUR tobit model are shown in figure A.4. Recall that only every 36th 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. Figures A.5 and A.6 show overlays of trace plots for the price parameters of the pooled and random effects models, respectively. 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.

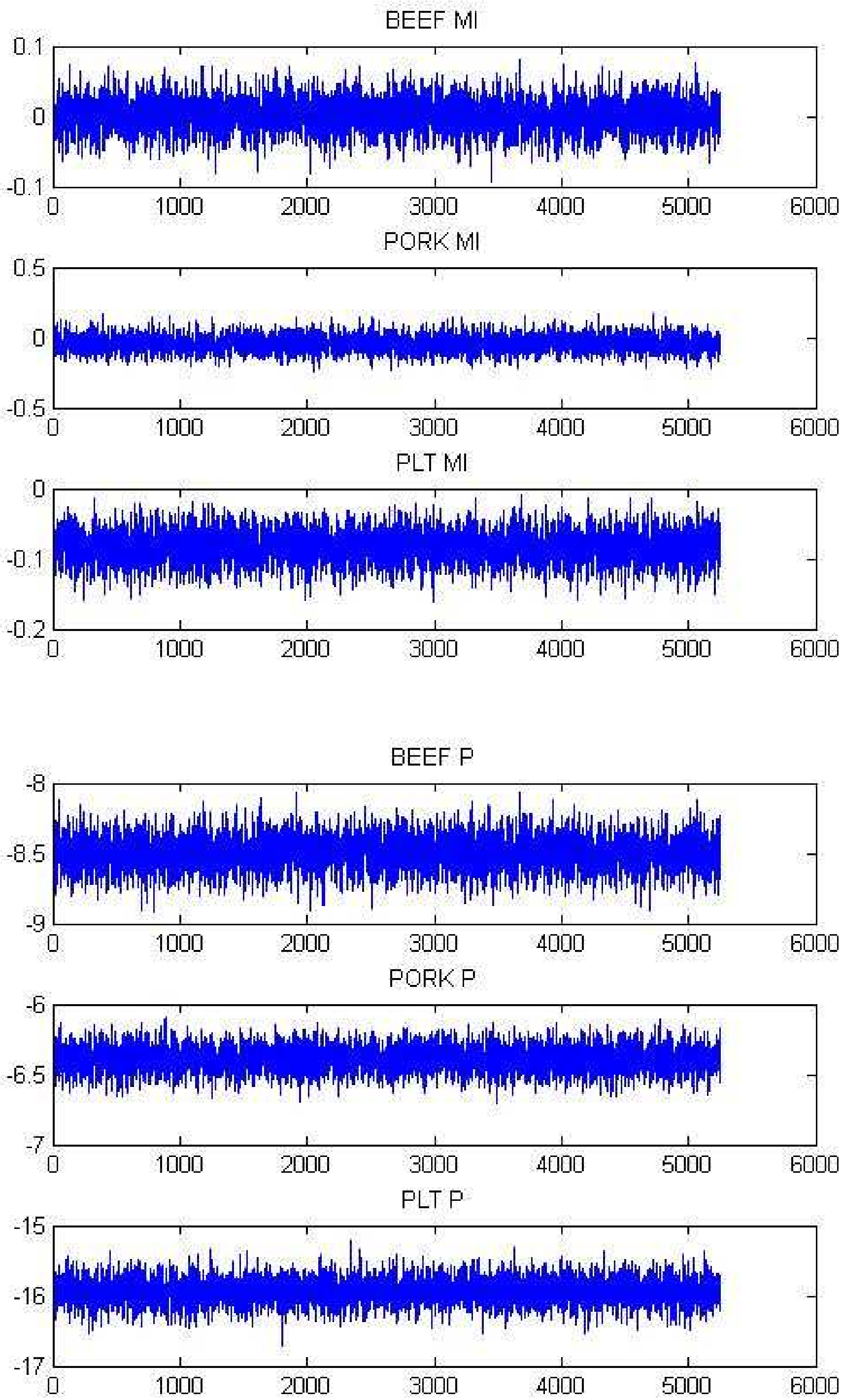

Figure A.1 Trace Plots of Media Index and Price Parameters from Pooled SUR Tobit Model

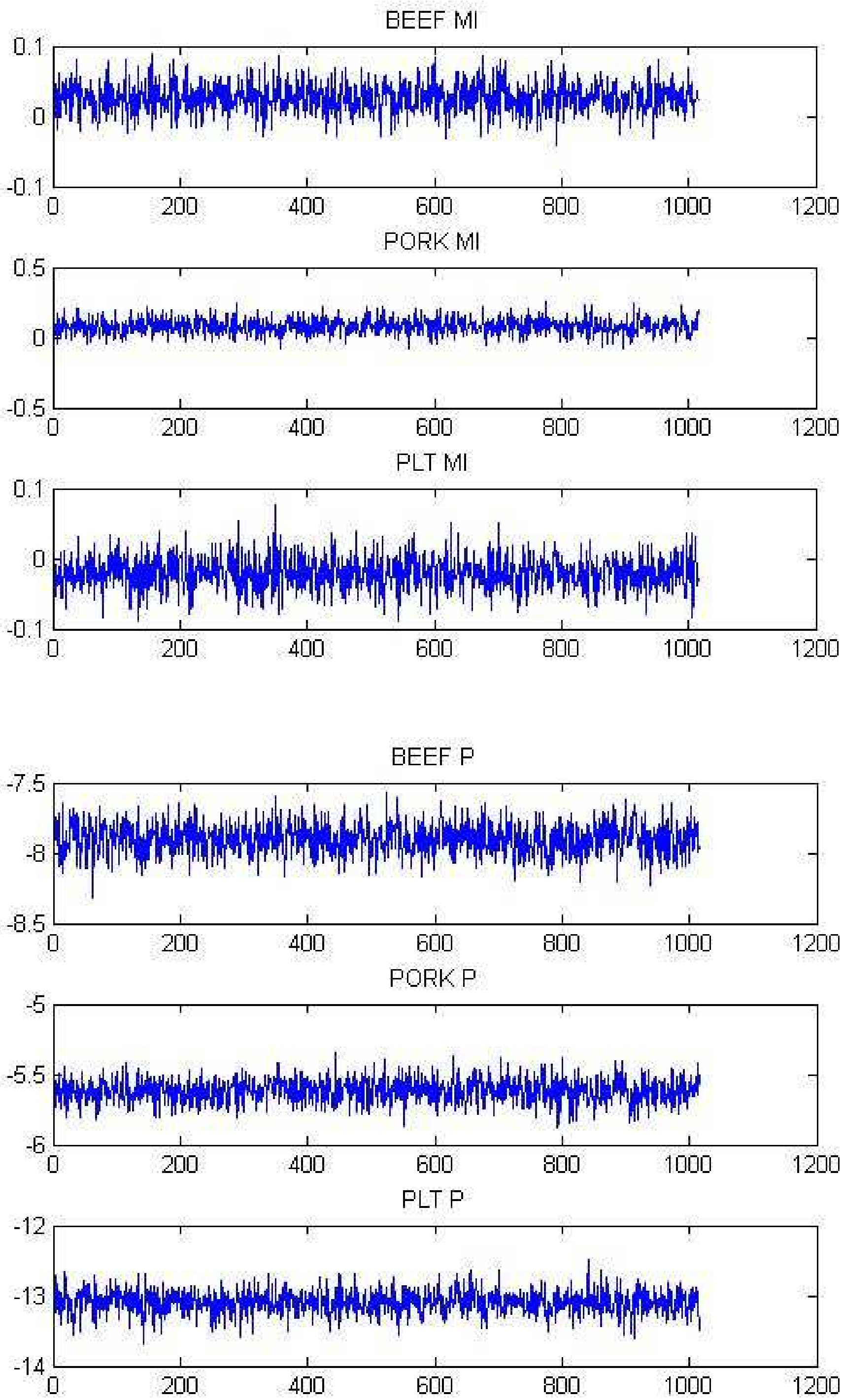

Figure A.2 Trace Plots of Media Index and Price Parameters from Random Effects SUR Tobit Model

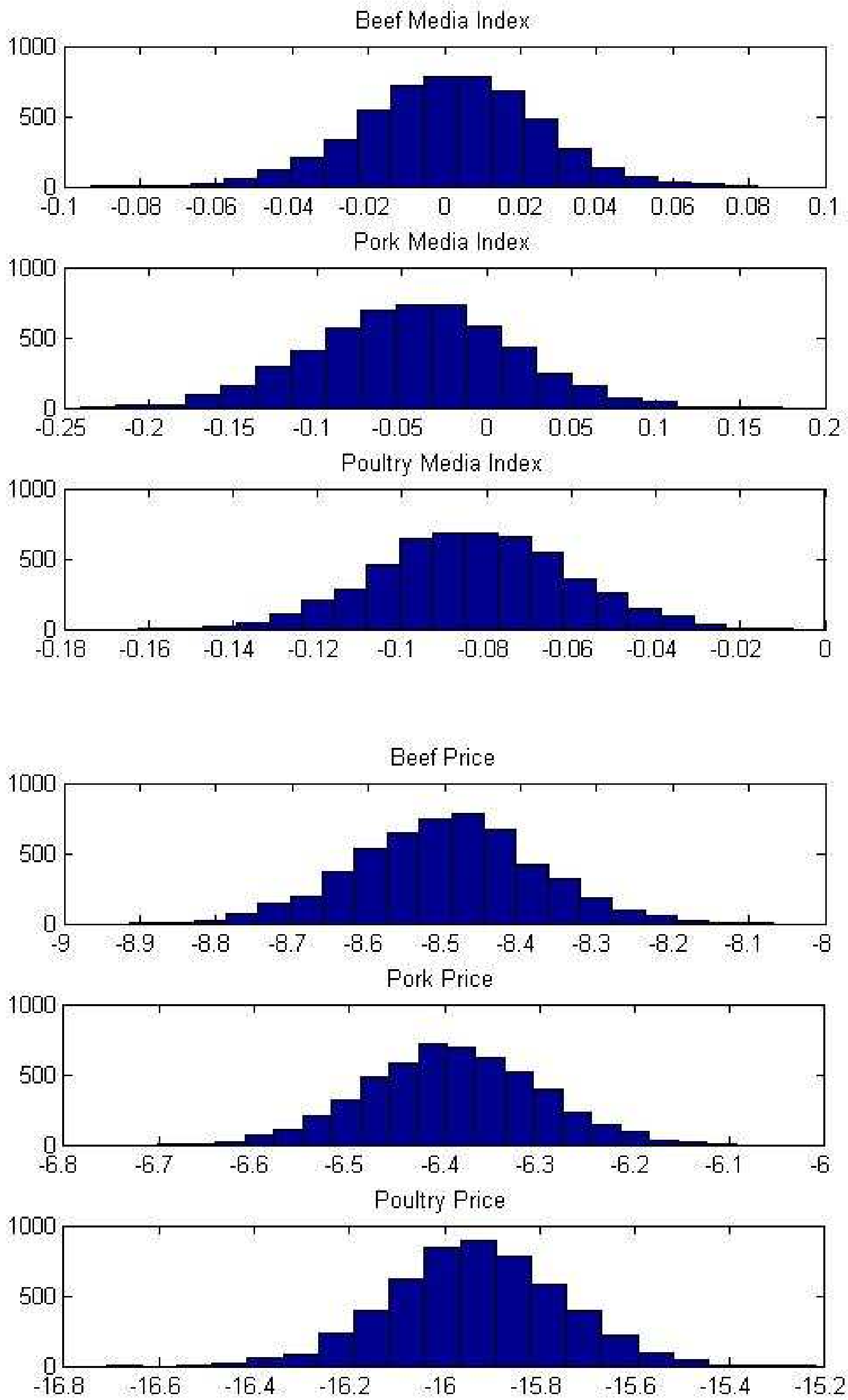

Figure A.3 Histograms of Media Index and Price Parameters from Pooled SUR Tobit Model

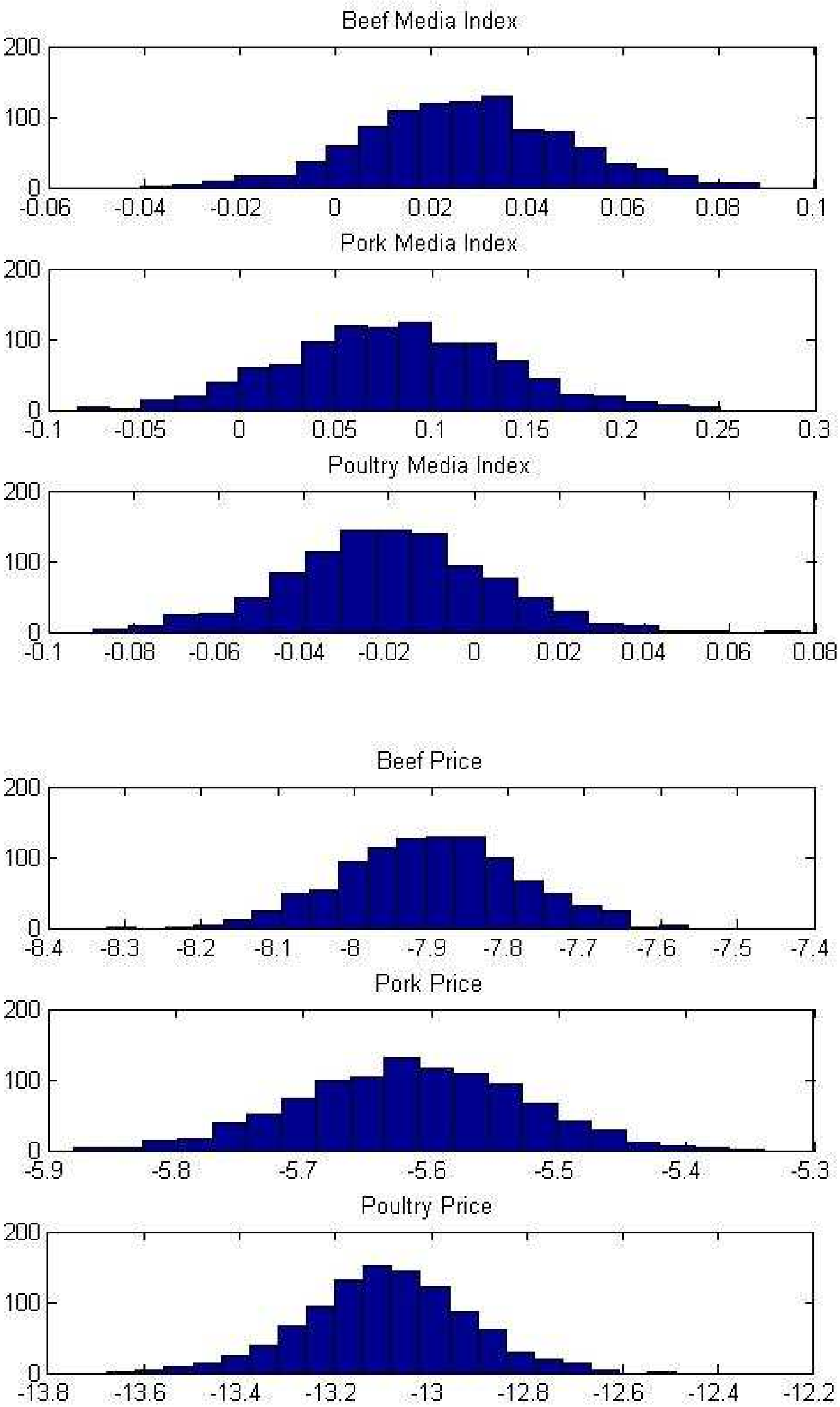

Figure A.4 Histograms of Media Index and Price Parameters from Random Effects SUR Tobit Model

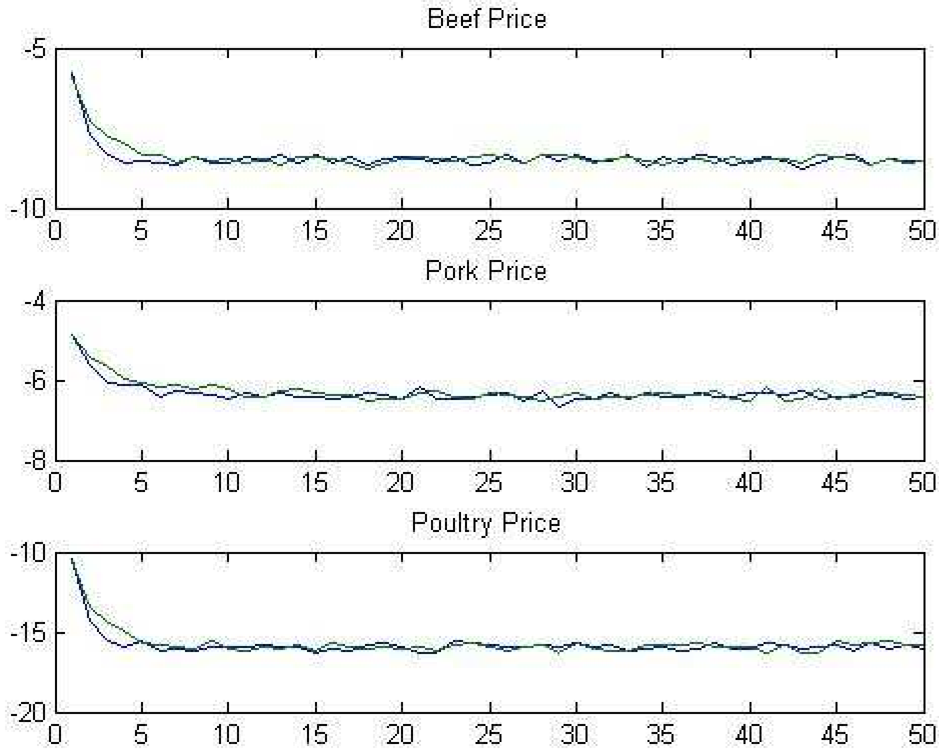

Figure A.5 Trace Plots of Price Parameters from Pooled SUR Tobit Model at Different Starting Values

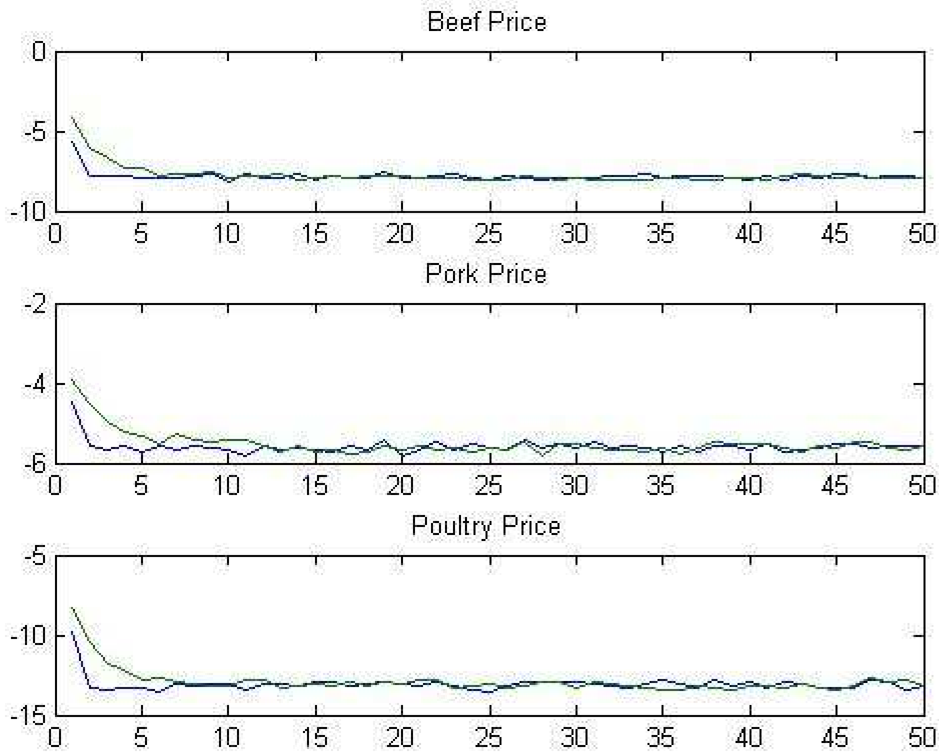

Figure A.6 Trace Plots of Price Parameters from Random Effects SUR Tobit Model at Different Starting Values