2006

Econometric essays on specification and estimation of demand systems

Anil Kumar Sulgham
Louisiana State University and Agricultural and Mechanical College, asulgh1@lsu.edu

ECONOMETRIC ESSAYS ON SPECIFICATION AND ESTIMATION OF DEMAND SYSTEMS

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in
The Department of Agricultural Economics & Agribusiness

by
Anil Kumar Sulgham
B.Sc., ANGRAU, 1998
M.S., University of Georgia, 2001
December 2006

To My Parents Vishwanath Reddy and Shyamala Devi

&

To My Beloved Wife Archana

ACKNOWLEDGMENTS

I wish to express my profound gratitude to Dr. Hector O. Zapata, my major professor, for his guidance and useful suggestions throughout the period of this research. I am grateful for his constructive criticisms and his comments on of this dissertation. Without his tireless assistance, leadership, and confidence in my abilities, this project would not have been possible. Working beside him was extremely demanding and challenging, and helped me significantly to get my best intellectual shape. Sincere thanks are extended to my graduate committee members, Dr. Roger Hinson, Dr. Wes Harrisson, Dr. Krishna Paudel, and Dr. Sudipta Sarangi for their constructive criticism and recommendations for this study.

Grateful acknowledgement is also expressed to the head and staff of the Department of Agricultural Economics and Agribusiness, for making it possible for me to do my research and studies here. Ms. Elizabeth Dufour did a lot to help me with proof reading and checking references. I am grateful for her assistance. I would also like to thank Mrs. Alicia Ryan for the hospitality during our stay. I am very fortunate for having had the chance to met excellent friends at LSU: Pramod Sambidi, James Henderson, Rebecca Mclaren, Tina Wilson, Brian Boever, Nandini Bandhopadhya, and Sachin Chintawar, for their invaluable friendship.

Finally, I am especially thankful to my family for their love and support. To my wife Archana Sulgham, I appreciate her patience, understanding, encouragement and faith in me. It was comforting to know that I had her unconditional love and support. Above all, I give credit to God who gave me strength throughout the duration of this program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation focuses on two research themes related to econometric estimation of linear almost ideal demand systems (LAIDS) for U.S. meats. The first theme addresses whether nonstationarity (unit-roots and cointegration) contributes to a dynamic specification of LAIDS models. The results of the effect of nonstationarity are reported in two case studies. The second theme explores the relationship between age and household size with budget shares to specify semiparametric LAIDS model. The results are reported in a third case study that compares parametric and semiparametric models estimates of price and expenditure elasticities.

The first case study conducts a comparative analysis of elasticity estimates from static and dynamic LAIDS models. Historical meat consumption data (1975:1-2002:4) for beef, pork and poultry products were used. Hylleberg et al. (1990) seasonal unit roots tests were conducted. Unit roots and cointegration analysis lead to the specification of an ECM of the Engle-Granger type for the LAIDS model. Marshallian and compensated elasticities were generated from the static and dynamic LAIDS models. The study found some model differences in elasticity estimates and rejected homogeneity in the dynamic model.

The second case study evaluates the forecasting performance of static and dynamic LAIDS models. Forecast evaluation was based on mean square error (MSE) criteria and recently developed MSE-tests. The study found ECM-LAIDS model performs uniformly better under all forecasting horizons for the beef equation. However, in the case of the pork equation the static model performed better in one-step-ahead and two-step-ahead forecasting horizons while the dynamic model was superior in the three-step-ahead and four-step-ahead forecasting horizons using MSE comparisons. In testing, only the two-steps ahead was superior for pork.

The third case study specifies a semiparametric LAIDS model that maintains the linearity assumption of prices and total expenditures and allows nonparametric effects of age and household size. 2003 U.S. Consumer Expenditure Survey data for four meat products (beef, pork, poultry and seafood) were used in the study. Model fit and elasticity estimates revealed negligible differences exist between parametric and semiparametric models.

CHAPTER 1

INTRODUCTION

The econometric specification of food demand systems has been a topic of extensive research interest. Over the past two decades, parametric models have dominated the empirical literature on this theme. Although economic theory is generally silent regarding the functional form of econometric models, applied demand analysis provides two utility-based approaches of generating demand systems (Theil & Clements, 1987). One approach applies classical economic optimization by specifying a utility function, an indirect utility function, or a cost function. Examples in this class of models include classical demand systems with quantity dependent equations, linear expenditure systems, budget share demand systems from translog indirect utility functions, and almost ideal demand systems (AIDS). A second approach is more mathematical and flexible; it generates demand equations by defining the total differential equation for each food product and, as opposed to the first approach, does not require the algebraic specification of utility or cost functions. Examples of demand systems generated from this approach include the Rotterdam model and the Working's model.

In addition to the demand systems generated from theory, there are various adaptations on the models that are used in estimating complete systems, group-of-food-products demand systems, cross-sectional data models, and time series panel data models. Examples include Huang and Haidacher's (1983) estimated complete food demand system for U.S. data, using a constrained maximum likelihood approach; Bharghava's (1991) estimated nutrient demand system for rural India, using a panel model; Karagiannis et al., (2000) estimated Greek meat demand system, using a time series model; and Nayga's (1996) study of the impact of household characteristics on away-from-home wine and beer weekly expenditures in the United States. All of the above models, and most of the published work on demand systems, fall into a class of models known as parametric. In any parametric model, the functional shape of the relationship is predetermined. The quality of the resulting estimator depends on the correctness of this specification. If the model is misspecified, then inferences and forecasts from such models are inadequate. Recent developments in econometrics provide a richer class of models (nonparametric and semiparametric models) with potential applications in the estimation of demand systems. Hence, the search for model structures that better fit theory and data is likely to continue.

One appeal for the application of semiparametric methods in the estimation of demand systems is their flexibility in capturing certain data patterns, such as nonlinearity, while keeping a parametric structure that may be suggested by economic theory. For example, if the focus of the econometric research is to estimate price and income elasticities, a semiparametric demand model can be specified with parametric-linear price and income effects, and with nonparametric demographic effects. Blundell et al., (1998) and Pendakur (1999) applied semiparametric models to estimate the nonlinear income and expenditure relationship in demand systems. It remains an empirical issue whether semiparametric specifications are an improvement to estimation of elasticities and forecasting with demand systems.

This dissertation focuses on two research themes related to econometric estimation of linear almost ideal demand systems (LAIDS) for U.S. meats. The first theme addresses whether nonstationarity (unit-roots and cointegration) contributes to a dynamic specification of LAIDS models. Consistent with the existing literature on demand systems, the results of the effect of nonstationarity are reported in two case studies. The second research theme explores the relationship between demographic factors (age and household size) and budget share to specify a semiparametric LAIDS model. The results are reported in a third case study that compares parametric and semiparametric models estimates of price and expenditure elasticities.

1.1 Problem Statement

The econometric specification of food demand systems has been of considerable research interest, due to the importance of elasticity estimates and commodity forecasts in marketing decisions and policy analysis. This dissertation addresses three research questions on the estimation of food demand systems. First, how do the nonstationary properties of time series data used in the estimation of demand systems affect elasticity estimates? Second, do dynamic demand systems improve out-of-sample forecasting performance? And third, is a semiparametric specification of food demand systems a more adequate approximation of cross-sectional data patterns? The analysis of these three questions will be reported via three econometric case studies.

1.2 Justification

The econometric model specification and estimation of demand systems has been a central theme in the analysis of U.S. meat consumption. Applied demand analyses are of interest because they provide updated estimates of price and income elasticities and demand forecasts. The quest for better and more reliable estimation methods is bound to continue. The first two of the questions relate to the common finding that most economic time series data tend to be nonstationary with a one-unit root. If a unit root exists in U.S. meat demand time series data, then there exists a possibility of cointegration, which would require the estimation of a vector error correction model (VECM). A VECM would make an elasticity estimate more reliable than those obtained from the usual least-squares based procedures. Also a VECM may improve demand forecasts. The third question is not related to time series, but deals with the specification of demand systems with micro data. Improved model specification in cross-sectional data may be obtained by combining linear parametric information from demand analysis with nonparametric smoothing around demographic information. For example, it is often reported that beef consumption increases with age and household size, but that the relationship may not be best represented by a smooth linear form between consumption (or budget shares) and age. Such behavior can be flexibly modeled through nonparametric techniques that smooth the demographic relationship between budget share and demographic variables while maintaining a parametric component that provides price and expenditure elasticity estimates. The approach is called a semiparametric method, and this study provides initial empirical evidence on the estimation of semiparametric demand systems.

1.3 Objectives

The general objective of this research is to assess the effect of dynamic and semiparametric estimation of an AIDS model for U.S. meats. The dissertation consists of three essays with the following specific objectives:

1) To compare elasticity estimates of static and dynamic AIDS model for U.S. meats,
2) To compare the predictive performance of static and dynamic AIDS model for U.S. meats,
3) To estimate a semiparametric AIDS model and compare elasiticity estimates to their parametric counterpart.

1.4 Data and Methodology

1.4.1 Objective 1

Until recently, the AIDS model has been estimated using primarily static models, ignoring the statistical properties of the data or the dynamic specification arising from time series analysis. The first case study on the dynamic specification and estimation of U.S. meat demand systems uses quarterly data over the period 1975(1)-2002(4) (a total of 112 observations). The quantity data are per capita disappearance data from the United States Department of Agriculture (USDA), Economic Research Service (ERS) supply and utilization tables for beef, pork, and poultry (sum of broiler, other-chicken, and turkey) gathered from online sources. The dynamic approach outlined by Karagiannis et al., (2000) is adopted. The time series properties of the data (unit-roots and cointegration) are used to establish a dynamic specification of linear AIDS (LAIDS) model. An error correction model (ECM) for LAIDS is established and econometrically estimated with an iterative seemingly unrelated regression (ISUR) procedure. We estimate both the static and dynamic models to compare tests on theoretical restrictions, such as homogeneity and symmetry conditions and elasticity estimates.

1.4.2 Objective 2

The second case study investigates forecast performance of the static and dynamic specifications of the LAIDS model for U.S. meat demand. Forecasts of domestic demand for U.S. meat products will be generated, using static and dynamic models. The forecast accuracy of the models will then be assessed. Accuracy measures are usually defined using forecast errors (i.e., the difference between the observed data and the forecast). Examples of such measures are the mean error (ME), the error variance (EV), the mean square error (MSE), and the mean absolute error (MAE), as defined by Diebold (1998). However, none of these approaches take into consideration the sample variability and uncertainty of the measures. Recent work revisited the concept of evaluating forecasts. West and McCracken (1999), for example, provided a comprehensive review on inference about a model’s ability to predict. The Diebold and Mariano test (1995) is used to compare the forecasting ability of the static and dynamic specifications of U.S. meat demand.

1.4.3 Objective 3

The final case study estimates the U.S. meat demand system with an emphasis on modeling the demographic characteristics using cross-sectional data extracted from the 2003 U.S. Consumer Expenditure Survey (CES). Demographic variables in a demand equation serve as shifters to permit more explanatory power to be achieved, allowing for a more accurate understanding of the behavior of consumers. Household size and age of reference persons are the two demographic variables of interest in this study. A popular LAIDS model framework is adopted for the estimation of the meat demand system.

Traditional parametric and flexible semiparametric techniques are used to estimate the demand system. One reason semiparametric procedures are of research interest lies in the periodical observation that cross-sectional consumption data patterns, particularly in relation to demographics, tend to behave nonlinearly. In the estimation of LAIDS models, these effects permeate throughout the system via parameter estimates that may not be the most accurate representation of the relationship between budget shares, prices, and expenditures. The semiparametric LAIDS model is specified with parametric LAIDS effects and nonparametric effects between budget shares and demographics. This research adds to that literature by conducting an empirical evaluation of model fit and elasticity estimates calculated from parametric and semiparametric models.

1.5 Overview of the Research

This research accomplishes the three objectives through a “journal-article-style” dissertation; each article is given in chapters three, four, and five. Chapter two includes a summary of consumer demand theory, and a review of previous work. The first case study, an estimation of U.S. meat demand systems using a dynamic approach, is presented in chapter three. Chapter four includes the second case study: investigating the predictive ability of static and dynamic demand systems. The final case study, comparing the price and income elasticities estimated from parametric and semiparametric models at disaggregate level, is presented in chapter five. The dissertation finishes in chapter six with conclusions.

CHAPTER 2

CONSUMER DEMAND THEORY AND LITERATURE REVIEW

The dissertation in general focuses on specification and estimation of food demand systems. The specific objectives of the research presented in chapter one, are addressed through three econometric case studies. The first two case studies concern the estimation of aggregate food demand systems using time series data and an analysis of forecast accuracy of the alternate specifications. Specifically, the first case study investigates the role of time series data properties in model specification and their influence on elasticity estimates. The second case study is an analysis of forecast accuracy of the dynamic specification resulting from the time series properties of the aggregate data. The final case study deals with a parametric and semiparametric estimation of the AIDS model, using household level cross-sectional data.

A fundamental framework of demand analysis is required to achieve the goals of this empirical work in applied economics. Also of importance is an understanding of prior research which will assist in proper placement of this work in the literature. The chapter is organized into two sections: A summary of consumer demand theory is presented in section (2.1) and a brief review of previous work on the specification and estimation of food demand systems is presented in section (2.2).

2.1 Consumer Demand Theory

Although economic theory is generally silent regarding the functional form of econometric models, applied demand analysis provides two utility-based approaches for generating demand systems (Theil & Clements, 1987). One approach applies classical economic optimization by specifying a utility function, an indirect utility function, or a cost function. Examples in this class of models include classical demand systems with quantity dependent equations, linear expenditure systems, budget share demand systems from translog indirect utility functions, and AIDS. A second approach is more mathematical and flexible; it generates demand equations by defining the total differential equation for each food product and, as opposed to the first approach, it does not require the algebraic specification of utility or cost functions. Examples of demand systems generated from this approach include the Rotterdam model and Workings model. This section provides a theoretical summary of the two approaches used to derive demand systems. Although the dissertation will not evaluate each of the various models mentioned below, the summary includes the derivation of the popular AIDS model and how it relates to other demand systems.

2.1.1 Utility Maximization

Consumer theory assumes that the most straightforward way to generate demand equations is to derive them by maximizing the utility function subject to the consumer's budget constraint. The utility framework is the foundation for index number theory, which includes the measurements of real income, the measurement of the effects of distortions such as commodity taxation, and the division of goods into groups that are closely related. In addition, the utility function generates the three major predictions of demand analysis: 1) the demand equations are homogeneous; 2) the substitution effects are symmetric; and 3) the substitution matrix is negative semidefinite. The utility function is denoted by

$$u = u(q_{1,\dots,}\, q_n) \tag{1.1}$$

where q_k is the quantity consumed of good k. The utility is maximized subject to a linear budget constraint

$$\sum_{k=1}^{n} p_k q_k = x\,, \quad k = 1,\dots, n, \tag{1.2}$$

where p_k is the price of the *k* good, and *x* is income or total expenditure. Theory assumes that the utility function is differentiable and there is nonsatiation, so that each marginal utility is positive

$$\frac{\partial u(q)}{\partial q_k} > 0 \quad k = 1,\ldots, n. \tag{1.3}$$

Mathematically, the consumer demand for a good derived from utility maximization is found by the Lagrangian method:

$$L(q,\lambda) = u = u(q_{1,\ldots,}\ q_n)\ +\ \lambda(x - \sum_{k=1}^{n} p_k q_k), \tag{1.4}$$

where λ is the Lagrangian multiplier interpreted as the marginal utility of income. Then, the first-order conditions of $\frac{\partial L}{\partial q_k}$, for k = 1, …, n, and $\frac{\partial L}{\partial \lambda}$ yield

$$\frac{\partial u(q)}{\partial q_k} = \lambda p_k\ , \tag{1.5}$$

$$x - \sum_{k=1}^{n} p_k q_k\ . \tag{1.6}$$

The first order conditions in equations (1.5) and (1.6) constitute *n+1* equations, which can be solved for the *n+1* unknowns q_1, …, q_n and λ. The resulting quantities are unique and positive for relevant values of prices and income. The optimal quantities depend on income and prices, so the demand functions may be written as

$$q_k = q_k(p_k, \ldots.., p_n, x) \quad k = 1, \ldots.., n. \tag{1.7}$$

The demand functions generated can be plugged back into the utility function to derive the indirect utility function given by:

$$u = u(q(x,p)) = u_I(x,p), \tag{1.8}$$

where *q* and *p* are vectors of *n* quantities and prices, and function $u_I(x,p)$ is the indirect utility function.

2.1.2 Indirect Utility Maximization

Indirect utility functions represent the maximum utility attainable corresponding to given values of prices and income. The theorem provided by Roy (1942) offers a second way to generate a system of demand equations from the indirect utility functions. Given an indirect utility function, Roy's identity,

$$q_k = -\frac{\partial u_I / \partial p_k}{\partial u_I / \partial x} \quad \text{k = 1, …., n,} \tag{1.9}$$

can be applied to generate the demand equations. Christensen et al., (1975) used this approach and introduced the translog indirect utility function to generate the translog demand system.

2.1.3 Cost Minimization and Consumer Demand

The consumer cost function is dual to the utility function in that it gives the minimum expenditure needed to reach a specified level of utility, when given the prices. The cost function is also referred to as the expenditure function and is expressed as a function of utility and price ($C(u, p) = C(u_I(x,p),p)$). The cost functions have a property

$$\frac{\partial C}{\partial p_k} = q_k \quad \text{k = 1, …., n,} \tag{1.10}$$

is referred to as the Shephard's lemma. Accordingly, a third approach to deriving demand equations is to specify the form of the cost function and then apply the Shephard's lemma. Deaton and Muellbauer (1980a, b) used this approach to generate the popular AIDS model.

2.1.4 Differential Demand Systems

In contrast to the above approaches to generate demand equations, the differential approach requires no algebraic specification of the utility function, the indirect utility function, or the cost function. The solution of a fundamental matrix equation is applied to derive a general system of differential demand equations. The total differential of equation (1.7) is

$$dq_k = \frac{\partial q_k}{\partial x} dx + \sum_{j=1}^{n} \frac{\partial q_k}{\partial p_j} dp_j \quad \text{k = 1, ….., n.} \tag{1.11}$$

Equation (1.11) is log transformed by multiplying both sides by p_k/x and using $w_k = p_k q_k / x$,

$$w_k d(\log q_k) = \frac{\partial p_k q_k}{\partial x} d(\log x) + \sum_{j=1}^{n} \frac{p_k p_j}{x} \frac{\partial q_k}{\partial p_j} d(\log p_j)\ . \tag{1.12}$$

Further simplification of equation (1.12), shown in Theil and Clements (1987), generates the demand equation for good *i* represented by

$$w_k d(\log q_k) = \theta_k d(\log Q) + \phi \sum_{j=1}^{n} \theta_{kj} d(\log \frac{p_j}{P^{`}})\ , \tag{1.13}$$

where $d(\log Q)$ is the Divisia volume index, and $d(\log P^{`})$ is the Frisch (1936) price index. Barten (1964) and Theil (1965) separately used the differential approach to generate the Rotterdam model.

2.1.5 Properties of Demand Functions

Deaton and Muellbauer (1993) reviewed the properties of consumer demand which provide reasonable restrictions to demand models. In many empirical works, these restrictions have been tested to confirm the theoretical validity of estimated demand functions. One of the most important properties of demand functions is adding up which is given by

$$\sum_{\mathrm{k}=1}^{n} p_{\mathrm{k}} h_{\mathrm{k}}(u, \mathrm{p}) = \sum_{\mathrm{k}=1}^{n} p_{\mathrm{k}} f_{\mathrm{k}}(\mathrm{p}, \mathrm{x}) = \mathrm{x}. \tag{1.14}$$

The estimated total value of both the Hicksian and Marshallian demands is total expenditures. In other words, the sum of the estimated expenditures on the different goods equals the consumer's total expenditures at any given time period. This property of demand provides a reasonable restriction, the so-called adding-up restriction (Deaton & Muellbauer, 1993). The adding-up restriction implies that

$$\sum_k p_k \frac{\partial q_k}{\partial x} = 1, \text{ equivalent to} \sum_k w_k e_k = 1, \tag{1.15}$$

where w_k is the budget share of good k and e_k is total expenditure elasticity. This implies that the marginal propensities to consume should sum to one. The second property of demand is homogeneity of degree zero in prices and total expenditures for uncompensated demand. If all prices and total expenditures are changed by an equal proportion, the quantity demanded must remain unchanged. This property is sometimes called the "absence of money illusion." The homogeneity property provides the homogeneity restriction which implies that, for k=1, . . ., n,

$$\sum_k p_k \frac{\partial q_i}{\partial p_k} + x\frac{\partial q_i}{\partial x} = 0, \text{ equivalent to} \sum_k e_{ik} + e_i = 0, \tag{1.16}$$

where $\sum_k e_{ik}$ is the sum of the own price elasticity and cross-price elasticities of the i^{th} good, and e_i is the total expenditure elasticity of the i^{th} good. The third property of demand is symmetry of the cross price derivatives of the Hicksian demands, that is,

$$\partial h_k(u,p)/\partial p_j = \partial h_j(u,p)/\partial p_k \text{ for all } i \neq j, \tag{1.17}$$

The symmetry expressed in equation (1.17) can be proven through Shephard's lemma (1953) and Young's theorem. Shephard's lemma is stated as:

$$h_k(u,p) = \partial c(u,p)/\partial p_k \text{ , } h_j(u,p) = \partial c(u,p)/\partial p_j$$

$$\partial h_k(u,p)/\partial p_j = \partial^2 c/\partial p_j\ p_k \text{, } \partial h_j(u,p)/\partial p_k = \partial^2 c/\partial p_j\ p_k \tag{1.18}$$

and in Young's theorem, $\partial^2 c/\partial p_j\ p_k$ equals $\partial^2 c/\partial p_j\ p_k$.

The last property of demand is negativity, which implies downward sloping compensated demand functions. Consumer demand theory has played an important role in the evolution of functional forms and econometric procedures used in the estimation of demand systems. A review of previous work on estimation of food demand systems presented in the following section provides insight into the role of theory in the specification of demand system.

2.2 Literature Review

Modern methods of estimating demand systems were initiated by Stone (1954). Individual equations for consumer goods were specified and estimated simultaneously; this led to a framework for simultaneously testing restrictions imposed by consumer theory (homogeneity and symmetry). Issues surrounding model specification and rejections of theoretical restrictions date back to these early efforts (e.g., Barten, 1969; Christensen et al., 1975; Deaton & Muellbauer, 1980). Barten (1969) rejected homogeneity based on the likelihood ratio statistic obtained from the maximum likelihood estimation of the Rotterdam model. Christensen et al., (1975) also concluded a rejection of homogeneity by using a transcendental logarithmic utility function to estimate the demand system. Deaton and Muellbauer (1980), who developed the AIDS model, rejected homogeneity based on F-tests. Deaton and Muellbauer assumed that the rejection of homogeneity is a symptom of dynamic misspecification.

Food demand has been of major interest in applied demand analysis over the past two decades. These studies can be classified into three categories: 1) cross-sectional, 2) time series, and 3) panel data. The focus of this research is estimation of food demand systems, using time series and cross sectional data for the U.S. population. Hence, the review focuses on empirical applications involving food demand in these two categories.

2.2.1 Cross Sectional Data Studies

Demand systems estimation makes use of household-level microdata, mainly to measure the effects of demographic variables. The estimation of demand systems using household-level data is more challenging than the conventional time-series data approach, for two reasons. First, for any given household, many of the goods have zero consumption, implying a censored dependent variable. Techniques which do not take this censored dependent variable into account will yield biased results. Second, household data are usually highly disaggregated across products, and it is next to impossible to estimate a completely disaggregated system because of the large number of products. Therefore, product aggregation is inevitable and is evident in the previous work. Cross-sectional studies for the U.S. consumer data include: Gao and Spreen (1994); Park et al., (1996); Byrne et al. (1996); Nayga (1996); Perali and Chavas (2000); Raper et al., (2002); Yen et al., (2002); Yen et al., (2003); and Dong et al., (2004).

Gao and Spreen (1994) estimated price and expenditure elasticities and the effect of household demographic variables on U.S. meat demand using the 1987-88 USDA household food consumption survey data. A hybrid demand system, which combines a modified generalized addilog system and a level version Rotterdam demand system, was developed and used as the analytical framework. The results suggested that region, ethnic background, household size, urbanization, food planner, health information, female household head employment status, and proportion of food expenditure on away-from-home consumption were the significant household characteristic and socio-economic variables. The finding supported speculation of other time-series meat demand studies, claiming both health concerns and convenience as the reasons for changes in consumer preference in favor of poultry and fish.

Nayga (1995) used the 1992 CES data to estimate the U.S. meat demand system. He adopted sample selection approach to estimate the demand system. The study found that beef and pork expenditures are positively related to household size. He also found that age is significantly related to expenditures on various meat products and expenditure on beef initially increases with age, and then declines.

Park et al., (1996) analyzed twelve food commodity groups according to household poverty status. They used the 1987-88 Nationwide Food Consumption Survey data. A Heckman two-step procedure for a system of equations was employed to account for bias introduced from zero expenditure on given commodities by a household. The second step of estimation involved the use of the linear expenditure system. Parameter estimates were used to obtain subsistence expenditures, own-price elasticities, expenditure elasticities, and income elasticities. Own-price elasticities were similar between the income groups for most commodities. However, income elasticities were consistently higher for the lower-income group.

Byrne and Capps (1996) used the two-step decision process for the estimation of the food-away-from-home demand system. The researchers estimated the demand system using a generalization of the Heien and Wessells (1990) approach. Household information gathered by the National Panel Diary Group was used for the analysis. Marginal effects were corrected by untangling the respective variable impacts on the inverse Mills ratio. Expenditure and participation probability elasticities were similar to previous studies. Income elasticities suggested that the food-away-from-home commodity is a necessary good for U.S. society.

Nayga (1996) studied the impact of household characteristics on away-from-home wine and beer weekly expenditures in the U.S. by applying the Heckman (1979) two-step procedure to the data extracted from the 1992 Consumer Expenditure Survey (CES). The study found that higher income households without children and headed by an older, white, and higher-educated individual spend more on wine away from home than do others.

Perali and Chavas (2000) developed an alternative econometric methodology to estimate a system of censored demand equations. The study used largely cross-sectional data drawn from Colombian urban households. The researchers used the two-step procedure for estimation. The first step used a Tobit model and introduced the methodology by specifying the AIDS model modified according to a translating and scaling demographic transformation. They then used the jackknife technique to estimate demand equations in unrestricted form and then recovered the demand parameters imposing the cross-equations restrictions by using minimum distance estimation.

Raper et al., (2002) analyzed food expenditures and subsistence quantities of poverty status and non-poverty status of U.S. households within a linear expenditure system that postulates subsistence quantities to be linear combinations of demographic variables. The data extracted from the 1992 CES were used in the study, applying the Heckman (1979) two-step procedure to estimate the demand system. The study presents analysis of expenditure elasticities, own-price elasticities, and subsistence quantities for each income group across nine broadly aggregated food commodity groups. The study found that elasticity estimates and subsistence quantity estimates differ across income groups.

Yen et al., (2002) estimated censored systems of household fat and oil demand equations with a two-step procedure, using cross-sectional data from the 1987-1988 U.S. Nationwide Food Consumption Survey. The study used a translog demand model and did not include the demographic characteristics. They found that own-price and total expenditure elasticities were close to unity, and compensated elasticities indicated net substitution among the products.

Yen et al., (2003) proposed a quasi-maximum-likelihood estimator and applied it to a censored translog demand system for foods. The research used food consumption by food stamp receiving households in the United States. Data are drawn from the National Food Stamp Program Survey (NFSPS). The study found that the procedure produces remarkably close parameter and elasticity estimates to those of the simulated-maximum-likelihood procedure.

Researchers also considered the two-step procedure, which produced different elasticities. Demands were found to be price elastic for pork and fish but price inelastic for all other food products, and the cross-price effects were less pronounced than own-price and total food expenditure effects.

Dong et al., (2004) extended the Amemiya-Tobin approach to demand system estimation using an AIDS specification. Under the Amemiya-Tobin approach, demand (share) equations are derived from a nonstochastic utility function and latent expenditures (shares) are hypothesized to differ from observed expenditures due to errors of maximization by the consumer, errors of measurement of the observed shares, or random disturbances that influence the consumer's decisions (Wales & Woodland, 1983). To account for these differences, error terms were added to the deterministic shares. The technique was applied to the 1998 expenditure survey data on Mexican households. They estimated twelve commodity demand models using simulated maximum likelihood procedures. Demographic characteristics such as household size, location, age, and number of children were also included in the model. The study found significant impacts of household size on demand elasticities.

A majority of earlier work used a parametric approach to estimate the effects of demographic characteristics. More recent interest has been on the application of semiparametric techniques using a single equation framework. Examples include Blundell et al., (1998) who used more flexible semiparametric models to estimate Engel curves for U.K. data. Similarly, Pendakur (1999) estimated semiparametric Engel curves using Canadian data. In this context, it is of interest to this dissertation to estimate a semiparametric AIDS model and compare generated elasiticities to their parametric counterparts.

2.2.2 Studies using Aggregate Time Series Data

Empirical analysis of food demand systems using aggregate time series data can again be divided into two sub-categories: a) static (majority of studies) and b) dynamic, based on the model specifications adopted in the estimation of the system.

The majority of the previous studies using U.S. data have adopted static models. Examples include: Eales and Unnevehr (1993); Moschini et al., (1994); Piggott (2003); and Dhar and Foltz (2005). Eales and Unnevehr (1993) developed the inverse AIDS (IAIDS) model in order to test the endogeneity of prices and quantities in the U.S. meat demand system. They found that IAIDS had all the desirable theoretical properties of the AIDS, except aggregation from the micro to the market level. The study employed annual data and found that both prices and quantities are endogenous within the entire meat market.

Moschini et al., (1994) derived a general elasticity representation of necessary and sufficient conditions for direct, weak separability of the utility function. The study used the Rotterdam model in the empirical analysis to test a few separable structures within a complete U.S. demand system, emphasizing food commodities. They found support for commonly used separability assumptions about food and meat demand.

Piggott (2003) introduced a new demand system, the Nested PIGLOG model, nesting thirteen other demand systems, including five that were also new. This new model and its nested special cases were applied to models of U.S. food demand that included food-at-home, food-away-from-home, and alcoholic beverages. The study found that although nested tests and out-of-sample forecasting performance favor generalizing models to a certain degree, statistically insignificant improvements to in-sample-fit and even poorer out-of-sample forecast accuracy undermine further generalizations. The study also found food-away-from-home to be price and income elastic, compared to food-at-home which also was price and income inelastic.

Dhar and Foltz (2005) used a quadratic AIDS model to estimate demand for various milk types. Scanner data from year 1997-2002 were used in the study. They studied the impact of labeling information (rBST-free and organic milk) and found that consumers are willing to pay significant premiums for such labels.

Previous studies using time series data and a dynamic specification to estimate the demand equations are grouped into a dynamic category. Examples include Pope et al., (1980); Chavas (1983); and Kastens et al., (1996); these researchers used a dynamic approach by including lagged variables or differencing approach, without formal testing for dynamic specification in the case of U.S. meat demand.

The first attempt at a dynamic approach dates back to the study of Pope et al., (1980) which used a flexible demand specification to test for homogeneity conditions and habit formation. The study applied Box-Cox transformations to four meat demand relations in order to allow for more flexible functional forms. The lagged terms were included in the model to measure habit persistence. Maximum likelihood techniques were used to estimate the parameters and homogeneity conditions, tested using likelihood ratio tests. The study rejected homogeneity, double log, and linear functional forms, based on the likelihood ratio tests.

Similarly, Chavas (1983) developed a method for investigating structural change in economic relationships in the context of a linear model. The approach assumes that the parameters can change randomly from one period to the next. The study applied the methodology to investigate the structural change in U.S. meat demand. The author identified structural changes that occurred in the 1970s for beef and poultry, but not for pork.

Kastens et al., (1996) estimated U.S. per capita food demand systems using an absolute price Rotterdam model, a first-differenced LAIDS and LAIDS model, and a first-differenced double-log demand system. The study used out-of-sample forecasting of annual U.S. per capita food consumption, applying data from 1923 to 1992 as a basis for model selection. They concluded that models with consumer theory, imposed through parametric restrictions, provided better forecasts than models with little theory-imposition; and the double-log demand system is a superior forecaster among alternate models.

One feature inclusive of the above studies is that the studies ignored formal testing procedures (unit-roots and cointegration tests) needed for establishing a dynamic specification. Studies that used unit-roots and cointegration tests to formulate dynamic specification to estimate food demand systems include Balcombe and Davis (1996); Karagiannis et al., (2000), and Fraser and Moosa (2002). Balcombe and Davis (1996) applied a LAIDS model to consumption in Bulgaria. They argued that the conventional estimation of the LAIDS should be done within the framework of contemporary time series methodology. The study applied canonical cointegrating regression procedure to estimate the demand system, and found that homogeneity and symmetry conditions hold in the case of the dynamic LAIDS model.

Karagiannis et al., (2000) presented a dynamic specification of the LAIDS based on recent developments in cointegration techniques and error correction models. The study used Greek meat consumption data over the period 1958-1993, and it was found that the proposed formulation performed well on both theoretical and statistical grounds, as the theoretical properties of homogeneity and symmetry were supported by the data. They also found beef and chicken to be luxuries while mutton-lamb and pork were necessities. All meat items were found to be substitutes to one another, except chicken and mutton-lamb, and pork and chicken.

Fraser and Moosa (2002) incorporated a stochastic trend and seasonality into the LAIDS model, using Harvey's structural time series methodology. They estimated the U.K. meat demand system, using three versions of the LAIDS model (deterministic trend and deterministic seasonality, stochastic trend and deterministic seasonality, and stochastic trend and seasonality). The study concluded that the structural time series model with stochastic trend and seasonality performed better in terms of model diagnostics, goodness-of-fit, and out-of-sample forecasting.

Studies by Ng (1995); Balcombe and Davis (1996); Attfield (1997); and Karagiannis et al., (2000) have suggested that when using time series models where the data have appropriate time-series properties (unit-roots and cointegration), one would find it likely that neither homogeneity nor symmetry is rejected. Ng (1995) concluded that homogeneity holds in many cases, using techniques including cointegration analysis. Attfield (1997) found that homogeneity holds by applying the triangular error correction procedure to the LAIDS model. Balcombe and Davis (1996) used the canonical cointegrating regression procedure for estimating the LAIDS. Karagiannis et al., (2000) outlined the potential use of an error correction model (ECM) of the LAIDS.

In the estimation of U.S. meat demand systems, often time series properties of data and the potential dynamic specification have been ignored. In the context of recent developments, the first case study of the dissertation focuses on an empirical analysis of U.S. food demand systems, using time series techniques. The role of time series properties of the data (unit-roots and cointegration) in the dynamic specification of an AIDS model is investigated, and the elasticity estimates generated from static and dynamic models are compared.

2.2.3 Forecasting Studies

Demand models are often used for forecasting, and forecast accuracy is of importance to forecast practitioners and followers. Out-of-sample forecasts have been used to measure forecast accuracy of the estimated demand systems in previous studies. Examples include Kastens et al., (1996); Chambers and Nowman (1997); Fraser and Moosa (2002); and Wang and Bessler (2003). Kastens et al., (1996) used out-of-sample forecasting of annual U.S. per capita food consumption, applying data from 1923 to 1992 as a basis for model selection. They used the root mean square error (RMSE) and mean absolute error (MAE) criteria to conclude that models with consumer theory imposed through parametric restrictions provide better forecasts than models with little theory-imposition, and the double-log demand system is a superior forecaster among alternate models.

Chamber and Nowman (1997) used the AIDS model as a representation of long run demands in both discrete time and continuous time error correction models. Out-of-sample forecasts were used to generate forecasts of budget shares beyond the sample period. The study used RMSE and MAE criteria to determine that continuous time adjustment mechanisms, based around fully modified estimates of the long run preference parameters, provide a remarkably accurate method of forecasting budget shares.

Fraser and Moosa (2002) estimated the U.K. meat demand system, using three versions of the LAIDS model (deterministic trend and deterministic seasonality, stochastic trend and deterministic seasonality, and stochastic trend and seasonality). The study used out-of-sample forecasts to perform the Ashley, Granger, and Schmalensee test (1980) for model selection.

Wang and Bessler (2003) estimated the U.S. meat demand system using quarterly data on meat. Researchers used Rotterdam model, static LAIDS model and vector error correction model

(VECM). The study used out-of-sample forecasts to perform the Diebold and Mariano test (1995) for model selection. Researchers concluded that VECM performed better in forecasting when compared to LAIDS and Rotterdam models.

The evaluation of forecasting accuracy via measures of point estimates is a well established practice in the forecasting literature. The mean error (ME), the error variance (EV), the mean square error (MSE), and the MAE are often used for evaluating forecasting performance. The usual practice for choosing among alternative forecasting models has been to select a model that shows a lower accuracy measure, but with no attempt in general to assess its sampling uncertainty. In this sense, the work by Parks (1990) is a good exception. More recently, the sampling uncertainty of point estimates of forecast accuracy has received considerable attention in econometric and forecasting literature (Diebold and Mariano, 1995; West, 1996; West and McCracken, 1998). This rich set of contributions allows for the evaluation of alternative forecasting demand models, one of the specific objectives of this research. The second case study examines out-of-sample forecast accuracy of two alternative specifications (static versus dynamic LAIDS) for the U.S. meat demand system, using the recently developed Diebold and Mariano (1995) test.

CHAPTER 3

A DYNAMIC AIDS MODEL FOR U.S. MEATS

3.1 Introduction

Static demand models express the relationship between budget shares, prices, and expenditures as a contemporaneous relationship. Early research in demand analysis was initiated by Stone (1954), who included a group of equations (one for each consumer good) in the system, and then estimating the equations simultaneously, there by adopting a static model. Since then, there have been numerous empirical studies of demand systems (Barten, 1969; Christensen et al., 1975; Deaton & Muellbauer, 1980) using static models. Deaton and Muellbauer (1980), who developed the almost ideal demand system (AIDS), were the first to acknowledge that the model suffers from dynamic misspecification.

Although the choice of an adequate functional form for demand systems remains a topic of empirical debate, the AIDS model has emerged as a popular functional form in empirical demand analysis. Until recently, the AIDS model has been estimated using a static approach, ignoring the statistical properties of the data or the dynamic specification arising from time series analysis. Recent developments in time series analysis offer new approaches to the dynamic specification of U.S. meat demand systems.

Recent studies (Ng, 1995; Attfield, 1997; Karagiannis & Mergos, 2002) have suggested that inconsistency between theory and data in demand analysis may be related to inappropriate modeling of time-series data. Ng (1995), using cointegration analysis, concluded that homogeneity holds in many cases. Attfield (1997) found that homogeneity holds when applying a triangular error correction procedure to almost ideal demand systems (AIDS). Balcombe and Davis (1996) proposed a canonical cointegrating regression procedure for estimating the AIDS model. This procedure is used in cases where prices follow a distributed lag process, or there is a seasonal pattern. Karagiannis et al., (2000) outlined the potential use of an error correction model (ECM) of the AIDS model.

The general approach followed is conditioned on the view that there may exist a long-run 'equilibrium' cointegrating demand system, measuring the long-run effects of prices and income on the demand for U.S. meats. New information and fluctuation in prices and income might disrupt the equilibrium and the process of adjustment may be incomplete in any single period of time. In the period before these adjustments are completed, consumers will be 'out of equilibrium,' and their short-run responses to changes in prices and income may provide little guide as to their long-run effects. In modeling the dynamics in meat consumption, we adopt a methodology for testing and setting up an error correction form of demand systems. The paper provides empirical evidence and measures of elasticity estimates of an ECM-AIDS for meat demand in the U.S. over the period 1975(1)-2002(4).

The rest of this paper is organized as follows. The empirical model is presented in section two. Data sources and descriptive statistics are summarized in section three. Section four presents estimation methodology. Empirical results and elasticity analysis are presented in five. Finally, the summary and conclusions are presented.

3.2 Empirical Model

The AIDS model has many desirable attributes: (a) it is an arbitrary first order approximation to any demand system; (b) it satisfies the axioms of choice in consumer theory; (c) it aggregates over consumers; and (d) it is easy to estimate. The estimated coefficients in a linear approximate almost ideal demand system (LAIDS) model are easy to interpret. It has been extensively used in empirical work (Green & Alston, 1990; Chalfant, 1987). Following past literature, meat is treated as a weakly separable group comprised of beef, pork, and poultry (chicken and turkey) in which consumption of an individual meat item depends only on the expenditure of the group, the prices of the goods within the group, and certain introduced demand shifters. The general specification of the AIDS model is given by:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(M / P) \tag{1}$$

where w_i is the expenditure share for the i^{th} good (i =beef, pork, and poultry), α_i is the constant coefficient in the i^{th} share equation, γ_{ij} is the slope coefficient associated with the j^{th} good in the i^{th} share equation, p_j is the price of the j^{th} good, and M is the total expenditure on the system of goods given by the following equation: $M = \sum_{i=1}^{n} p_i q_i$ where: q_i is the quantity demanded for the i^{th} good. P is a general price index defined by:

$$\log P = \alpha_0 + \sum_{i=1}^{n} \alpha_i \log p_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \log p_i \log p_j \,. \tag{2}$$

To comply with the theoretical properties of consumer theory, the following restrictions are imposed on the parameters in the AIDS model:

- *Adding up restriction*: $\sum_{i=1}^{n} \alpha_i = 1, \sum_{i=1}^{n} \beta_i = 0, \sum_{i=1}^{n} \gamma_{ij} = 0$, allowing the budget share to sum to unity,
- *Homogeneity*: $\sum_{j=1}^{n} \gamma_{ij} = 0$, which is based on the assumption that a proportional change in all prices and expenditure does not affect the quantities purchased. In other words, the consumer does not exhibit money illusion,
- *Symmetry*: $\gamma_{ij} = \gamma_{ji}$, represents consistency of consumer choices.

In empirical studies, to avoid non-linearity and reduce multi-colinearity effects in the model, equation (2) is sometimes approximated by a Stone index defined as $\log P = \sum_{i=1}^{n} w_i \log p_i$. We use the simple linear AIDS (LAIDS) model in our empirical investigation. Researchers are mostly interested in the demand elasticities, which are easy to estimate in this flexible functional form of the LAIDS model. According to Green and Alston (1990), elasticities in LAIDS can be expressed as: $\eta_i = 1 + \beta_i / w_i$ for income elasticity and $\eta_{ij}^* = -\delta_{ij} + w_j + \gamma_{ij} / w_i$ for compensated elasticities. The uncompensated elasticities are computed from $\eta_{ii} = -\delta_{ij} - \beta_i + \gamma_{ij} / w_i$, and $\eta_{ij} = \gamma_{ij} / w_i - \beta_i \, w_j / w_i$.

The static LAIDS model has come to be known as the long run LAIDS model (Duffy, 2003). The long run model implicitly assumes that there is no difference between consumers' short run and long run behavior, that is, the consumers' behavior is always in "equilibrium." However, in reality, habit persistence, adjustment costs, imperfect information, incorrect expectations, and misinterpreted real price changes often prevent consumers from adjusting their expenditure instantly to price and income changes (Anderson and Blundell, 1983). Therefore, until full adjustment takes place, consumers are "out of equilibrium." It is therefore necessary to augment the long-run equilibrium relationship with a short-run adjustment mechanism. It is well known that most economic data are nonstationary, and the presence of unit roots may invalidate the asymptotic distribution of the estimators. Therefore, traditional statistics such as t, F, and R-square are unreliable, and least squares estimation of the static LAIDS tends to be spurious.

The concepts of cointegration and error correction models (ECM) were first proposed by Engle and Granger (1987) and have been widely used by researchers and practitioners in modeling and forecasting macroeconomic activities over the last decade. Engle and Granger

(1987) showed that the long-run equilibrium relationship can be conveniently examined using the cointegration technique, and the ECM describes the short-run dynamic characteristics of economic activities. By transforming the cointegration regression into an ECM, both the long-run equilibrium relationship and short-run dynamics can be examined. Secondly, the spurious regression problem will not occur if the variables in the regression are cointegrated.

The variables in equation (1) (budget shares, log of prices and total expenditures) must be tested for unit roots before examining if cointegration exists. The Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1981), Phillips-Perron (PP) (Phillips, 1987; Perron, 1988) statistics are commonly used for testing unit roots. In agricultural economics, meat markets time series exhibit substantial seasonality, therefore, there is a possibility that there may be unit roots at seasonal frequencies. Hence, there exists a need for testing procedures that accounts for seasonality such as Hylleberg et al. (1990) seasonal unit roots test. Once the orders of integration of the variables have been identified, either the Engle and Granger (1987) two-stage approach or the Johansen (1988) maximum likelihood approach can be used to test for the cointegration relationship among the variables in the models.

If cointegration is found, then the LAIDS model is estimated as an ECM. Applications of the ECM-LAIDS can be seen in the studies of demand for food, and meat products (Balcombe & Davis, 1996; Attfield, 1997; Karagiannis et al., 2000; Karagiannis & Mergos, 2002). The ECM of the LAIDS model used in this paper follows Karagiannis and Mergos (2002) and is given by

$$\Delta w_i = \sum_{k=1}^{l} \delta_k \Delta w_{t-k} + \sum_{j=1}^{n} \gamma_{ij} \Delta \ln p_j + \beta_i \Delta \ln(M / P) + \lambda_i \mu_{it-1} + \mu_t \qquad (3)$$

where Δ refers to the difference operator and μ_{it-1} includes lagged residuals from the first step OLS regression, which measures the feedback effects. The parameter λ_i is the error-correction term, being the deviation of actual budget shares in the previous period, w_{t-1}, from the values that were desired on the basis of the information then available, w_{t-1}^{*} (where the asterisk denotes a desired value). Consumers in the current period attempt to change w_t from its value in the previous period, w_{t-1}, with the goal of closing some of the gap that may have existed between w_{t-1} and w_{t-1}^{*}. These adjustments move budget shares in the direction of their desired values, eventually establishing long-run equilibrium. Lagged budget share effects are likely to be important in the demand for meat products, because of the influence of habit formation. The parameters δ_k and λ_i are to be estimated. Theoretical restrictions adding-up, homogeneity, and symmetry discussed earlier are also applicable to the ECM-LAIDS model.

3.2.1 Error Correction Modeling

Engle-Granger (1987), two-step methodology is used to estimate the ECM-LAIDS model. The Engle-Granger two-step method proceeds as follows. In the first step, tests of unit roots are applied to budget shares, prices and expenditures. If unit roots are found, a cointegrating regression of budget shares on prices and total expenditures is estimated and diagnostic tests of cointegration are applied to this last regression. If cointegration is found, the lagged residuals from the cointegrating regression are included as an independent variable in the specification of an ECM-LAIDS.

The second step involves seemingly unrelated regression on the ECM-LAIDS model represented in Eq. 3. A system of equations can be applied using the differenced variables and the residuals obtained from the first step to account for unit roots and cointegration. The system of equations in the matrix notation is given by

$$\begin{bmatrix} y_b \\ y_p \\ y_c \end{bmatrix} = \begin{bmatrix} x_b & 0 & 0 \\ 0 & x_p & 0 \\ 0 & 0 & x_c \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}. \tag{4}$$

The left hand side variables y_i in the Eq. 4 are the budget shares expressed in first differences, and the right hand side variable x_i includes the lags of first differenced budget shares, accounting for habit persistence, first differenced prices, first differenced expenditure, and lagged residuals from the first step OLS regression. The system of equation is then estimated, using the seemingly unrelated regression procedure.

The following assumptions are made for the equation errors ε_1, ε_2, and ε_3:

1. All errors have zero mean: E(ε_i) = 0; for i = 1,2,3;
2. In a given equation, the error variance is constant over time, but each equation can have a different variance.
3. Two errors in different equation, but corresponding to the same time, are correlated (contemporaneous correlation).
4. Errors in different time periods are not correlated.

3.3 Data

The data used in the analysis are quarterly observations over the period 1975(1)-2002(4), providing a total of 112 observations. The quantity data are per capita disappearance data from the United States Department of Agriculture (USDA), Economic Research Service (ERS) supply and utilization tables for beef, pork, and poultry (sum of broiler, other-chicken, and turkey), gathered from online sources. The beef price is the average retail choice beef price, the pork price is the average retail pork price, and the poultry price was calculated by summing quarterly expenditures on chicken, using the average retail price for whole fryers, and quarterly expenditures on turkey, using the average retail price of whole frozen birds, divided by the sum of quarterly per capita disappearance on chicken and turkey, similar to Piggott and Marsh, (2004). The retail prices are deflated and used for further analysis, using the consumer price index for nonfood items (base year being 1984). The total expenditures on meat and budget shares of each meat product are estimated, using the price and quantity information discussed above. Table 3.1 provides descriptive statistics of quarterly data used in this study.

The descriptive statistics indicate that the per-capita mean consumption of beef was 18.52 pounds, with mean share equivalent to 54%. The per-capita mean consumption of pork was 12.75 pounds, with mean share equal to 27%, and the mean consumption of poultry was 18.64, pounds with mean share around 19%. The variance measure indicates that the beef expenditure share exhibits the highest variability followed by poultry and pork expenditure shares. In the case of prices, the average price of beef was highest at 2.52 $/lb followed by pork (1.98 $/lb) and poultry (0.86 $/lb). The standard deviation measure for prices indicates that beef and pork prices exhibit higher variability when compared to poultry prices. The mean total expenditure on meat was around $ 87.64.

Table 3.1. Descriptive Statistics of Beef, Pork, and Poultry Demand in the U.S, 1975(1)-2002(4).

Variable	Mean	Std. Dev	Minimum	Maximum
Beef consumption (lbs/capita)	18.529	2.111	15.891	24.465
Pork consumption (lbs/capita)	12.753	0.939	9.658	14.863
Poultry consumption (lbs/capita)	18.643	4.132	10.335	26.157
Retail beef price ($/lb)	2.516	0.501	1.348	3.448
Retail pork price ($/lb)	1.983	0.401	1.208	2.749
Retail poultry price ($/lb)	0.864	0.139	0.601	1.111
Meat Expenditure($/capita)	87.643	16.306	50.289	119.14
Beef share	0.543	0.042	0.434	0.615
Pork share	0.272	0.019	0.253	0.322
Poultry share	0.185	0.036	0.123	0.242

3.4 Estimation Methodology

The data used in the study were seasonally unadjusted quarterly observations. Hence, seasonal unit roots tests suggested by Hylleberg et al., (1990) were applied to each series. Using Osborn et al. (1988) notation I(a, b), the first argument (a) representing the non-seasonal (first) differencing, and the second argument (b) representing the order of seasonal differencing necessary for stationarity. Thus, a quarterly series is said to be I(1, 1) if it requires both one quarter and seasonal (four quarter) differencing to become stationary. An I(0, 1) series requires only seasonal differencing; an I(1, 0) series needs only one quarter differencing; and an I(0, 0) series is stationary in levels and needs no differencing.

Test results are presented in Table 2. The null hypotheses for these tests states that the series investigated are an I(0, 1). The tests are based on the following regression after augmentation with lagged dependent variables and deterministic components:

$$y_{4t} = \pi_1 y_{t-1} + \pi_2 y_{2t-1} + \pi_3 y_{3t-2} + \pi_4 y_{3t-1} \qquad (5)$$

where $y_{1t} = (1 + L + L^2 + L^3)x_t$, $y_{2t} = -(1 - L + L^2 - L^3)x_t$, $y_{3t} = -(1 - L^2)x_t$, and $y_{4t} = (1 - L^4)x_t$. Equation (4) is estimated initially with all lagged values of the dependent variable up to a maximum lag of eight quarters, plus a constant, trend, and three seasonal dummies. A testing down procedure is then followed to eliminate insignificant lagged values of the dependent variable, working from the longest lags towards the shortest, but always subject to the condition that the residuals exhibited no evidence of serial correlation up to the fourth order (Duffy, 2003).

The null hypothesis, that X_t is I(0, 1), is not rejected if all $\pi_i = 0$ (i = 1, 2, 3, 4). This is tested by a joint F statistic, denoted as F_{1234} in Table 2. The alternative hypotheses that are worth considering are that each variable is I(1, 0) or I(0, 0). An insignificant t-value for π_1, combined with a significant F_{234} statistic, implies that the series is I(1, 0), whereas a significant t-statistic for π_1 and a significant F_{234} statistic indicates that the series is I(0, 0). The F_{1234} statistics in Table 3.2 indicate that all of the series used in this study are not I(0, 1). The combination of insignificant t-ratios for π_1 (implying non-rejection of $\pi_1 = 0$) and significant values for F_{234} (rejecting the presence of unit roots at the seasonal frequency) leads to the conclusion that the all the series are I(1, 0). Therefore, any remaining seasonality in the series would be deterministic and can be modeled with, for instance, seasonal dummy variables.

Table 3.2. Seasonal Unit Root Test Results for Budget Share, Prices, and Total Expenditure (Hylleberg et al., 1990).

Variable	*t*-statistic for Π_1	F_{234}	F_{1234}	Augmentation of Lags	Conclusion
W_1	-1.16	35.07	26.89	0	I(1,0)
W_2	-2.21	46.91	36.66	0	I(1,0)
W_3	-0.67	40.42	30.50	0	I(1,0)
ln p_1	-1.87	182.04	140.33	0	I(1,0)
ln p_2	-2.56	52.85	49.21	1	I(1,0)
ln p_3	-2.37	101.47	85.92	0	I(1,0)
ln (m/P)	-1.62	55.68	42.36	0	I(1,0)
Critical values (5%)	-3.53	5.99	6.47		

Notes: [a] Subscripts refer to (1) Beef, (2) Pork, and (3) Poultry.
The 5 % critical values are taken from Ghysels et al. (1994); they are appropriate for a test regression that includes constant, seasonal dummies, a linear trend and for that which is estimated from a sample size of 100 observations.

Fig. 3.1 and Fig. 3.2 show the time path of levels and the first differences of budget shares, expenditure, and price series, respectively. The time plots (Fig. 3.1) show that beef expenditure shares have a clear, seasonal pattern and distinctive, downward trend. The time plot for pork budget shares shows a seasonal pattern with small variation in early 1980, and then an increasing trend in a later part of the sample period. The poultry expenditure shares also exhibit a seasonal pattern with a distinctive increasing trend until year 2000, and then stable behavior. A distinctive feature in these time plots is that the expenditure shares for all categories appear to be nonstationary in levels. The time plots of first differences for all the budget shares appear to be stationary, consistent with the finding of the seasonal unit root test. The first differenced time plot for poultry shows very little volatility after the year 2000, thus making it simple to predict.

The time plots (Fig. 3.2) for a log of pork and poultry prices show seasonal behavior and a clearly increasing trend. The time plot for beef prices shows an initial increase, followed by stabilization in the early 1980s, and then an increasing trend from the year 1987 onward. The time plot of total expenditure exhibits seasonal behavior and a distinctively increasing trend after the year 1990. Time plots of the first differenced series for pork and poultry exhibit a high volatility initially, but taper toward the end of the sample period, making prediction simple. A distinctive feature of all the first differenced series is that they all appear to be stationary, showing a consistency with the findings of the seasonal unit root test.

Having established that the series are I(1,0) (each series contains a unit-root) we proceed to test for cointegration between budget shares, prices, and expenditures using Engle and Granger (1987) methodology. This method is based on testing whether ordinary-least squares (OLS) residuals from the cointegrating regression are stationary for each share equation. If the residuals are stationary, then there exists a cointegrating relationship. In the results from the ADF and PP tests reported in Table 3.3, only the residuals from budget-shares of poultry equation are nonstationary, and hence not cointegrated at the 5% significance level. The static cointegration tests are often considered to be low in power, while discriminating the alternative hypotheses. Banerjee et al. (1986) and Kremers et al. (1992) recommended a robust, dynamic, modeling procedure to perform a cointegration test. According to this methodology, an ECM is formulated (Eq. 3) and estimated. Then, the hypothesis that the coefficient of error correction term is not statistically different from zero is tested using a traditional *t*-test. If we fail to reject null hypothesis, the series concerned are not cointegrated. The residuals from the earlier cointegration regression are used as the ECM term in this step. Based on the statistical significance of λ_i, parameters associated with the ECM term, (Table 3.3) we conclude the existence of a cointegrated regression equation for all budget shares.

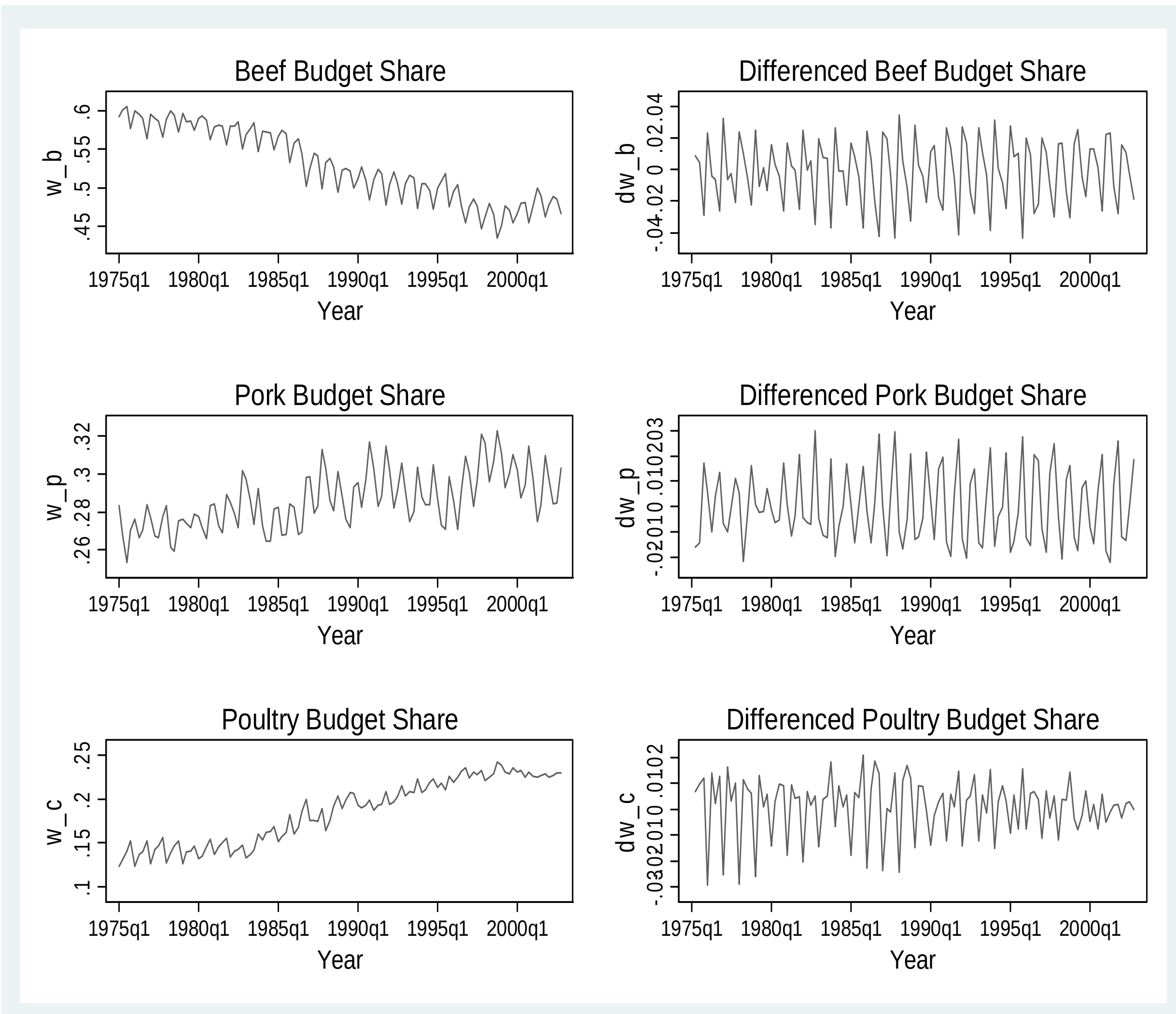

Figure 3.1. Time Plots in Levels and Differences of Budget Shares.

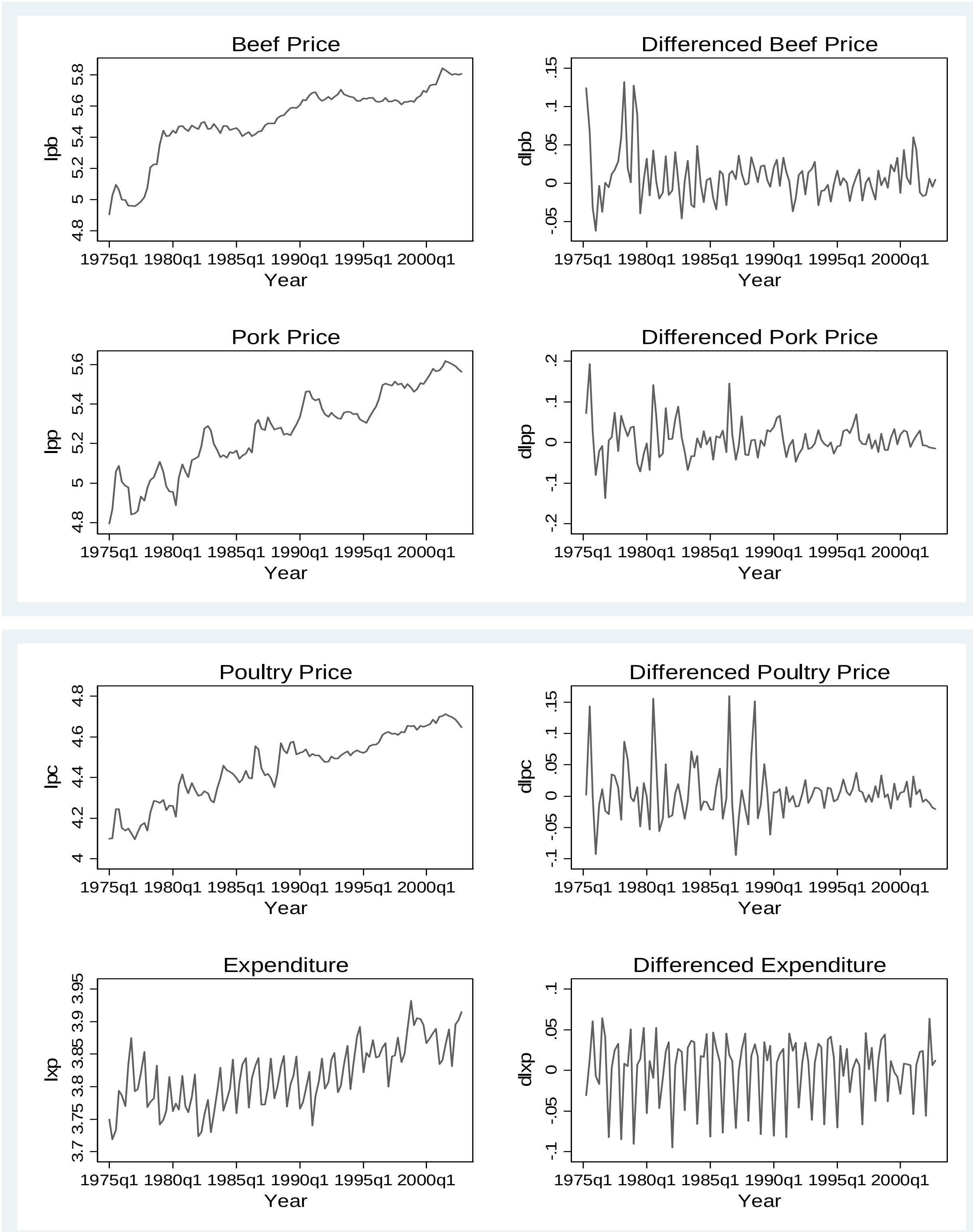

Figure 3.2. Time Plots in Levels and Differences of Prices and Expenditure.

Table 3.3. Static and Dynamic Cointegration Tests Results.

Equation	CI test[b]		Dynamic CI test[c]	
	ADF	PP	λ	t-value
W_1	-5.61	-7.294	-0.425	-5.09
W_2	-8.24	-8.315	-0.932	-8.69
W_3	-4.25	-4.595	-0.121	-2.51

Notes: Cointegration tests are based on regression including a constant term and a time trend.

[b]For the Engle Granger CI test, the tabulated critical value at 5% is 4.87.

[c]Based on estimation of Eq. (2).

Since the sum of all expenditure shares in the LAIDS model is equal to unity, the residuals variance-covariance matrix is singular. The usual solution is to delete an equation from the system and estimate the remaining equations, and then calculate the parameters in the deleted equation in accordance with the adding-up restrictions. In our case, we arbitrarily drop the poultry equation from the system. First, we estimate the unrestricted static LAIDS models using Eq. (1). We add the deterministic components in the form of seasonal dummies and a linear time trend in the model. Estimation is carried out implementing the maximum likelihood (ML) routines for seemingly unrelated regression (SUR). Later we impose the homogeneity and symmetry conditions separately and then combine them to estimate the restricted models. The likelihood ratios estimated from the unrestricted and restricted models are presented in Table 3.4. Results (p-values of > 0.05) indicate both homogeneity and symmetry conditions are satisfied by the static model.

3.5 Empirical Results

The estimates from the restricted, static LAIDS model are presented in Table 3.5 (homogeneity and symmetry constraints imposed). The parameter estimates of the beef budget share equation indicate a positive relationship with respect to its own price (0.066) and total expenditure (0.089), while it has a negative relationship with pork (-0.002) and poultry (-0.064) prices. The parameter estimates for pork share equation indicate a positive own-price coefficient (0.048), while the poultry price (-0.046) and total expenditure (-0.009) had negative coefficients. The parameter estimates for the poultry equation were derived by use of adding up constraints; the own-price coefficient (0.110) was positive, and the total expenditure had a negative coefficient (-0.079). The trend variable (time) was positive for pork and poultry share equations and negative for beef equations, confirming the initial graphical evidence. All the parameter estimates associated with the quarterly dummies were positive and significant for beef equations, while the first two quarters had a negative effect for pork and poultry equations.

Table 3.4. Results of Likelihood Ratio Tests for Theoretical Restrictions.

	Calculated x^2	p-Value
	Static-LAIDS	
Symmetry	0.72	0.3955
Homogeneity	0.06	0.9723
Homogeneity and Symmetry	0.84	0.8397
	ECM-LAIDS	
Symmetry	0.12	0.7245
Homogeneity	21.87	0.0000
Homogeneity and Symmetry	23.26	0.0001

With regard to the dynamic ECM-LAIDS model where the linear time trend is omitted, four lagged budget share variables were included as measures of habit persistence (Akaike information criteria was used to determine the number of lags). The Engle and Granger (1987) two-step approach is employed for estimating cointegrating regressions. The residuals from these regressions are obtained and incorporated into Eq. 3, and then the unrestricted ECM-LAIDS is estimated using the MLE of SUR procedure. The estimates are shown in Tables 3.6. The majority of the estimated parameters (δ_k) in the ECM-LAIDS are significantly different from zero, suggesting strong habit persistence. In other words, current period consumption of meat is influenced by previous period meat consumption. The coefficients of the error correction terms are all statistically significant at the 1% level, suggesting that any deviations of meat-spending from the long-run equilibrium are accounted for in the dynamic LAIDS model. The negative coefficients of the error correction terms for beef and pork (-0.133 and -0.291, respectively) suggest that deviations in previous period result in reduced budget shares. These adjustments move budget shares in the direction of their desired values, eventually establishing long-run equilibrium. Similarly, a positive coefficient for poultry (0.425) suggests that shocks in previous period leads increased budget shares in the current period. With regard to the restriction tests (see Table 3.4), unfortunately the ECM-LAIDS passes only the symmetry test at the 5% level, while failing the homogeneity test and the joint tests for both homogeneity and symmetry. The estimates from the restricted dynamic LAIDS model are presented in Table 3.6 (imposing both symmetry and homogeneity). The parameter estimates from both the restricted models (static and ECM) are used for elasticity analysis.

3.5.1 Elasticity Analysis

The estimated Marshallian elasticities from the static model are presented in the upper half of Table 3.7. The estimated own price elasticities were all negative, consistent with the demand theory suggesting a downward-sloping demand curve (-0.905, -0.818, and -0.191 for beef, pork, and poultry, respectively). These results mean that per capita beef consumption that is conditional on meat expenditure is more sensitive to its own-price change, while poultry consumption is least sensitive to changes in its own-price. The Marshallian own-price elasticities obtained from the ECM-LAIDS model are -0.53, -0.73, and -0.50 for beef, pork, and poultry, respectively, presented in the lower half of Table 3.7, suggesting pork consumption is more sensitive to its own price change in the short-run, while the effect of a price change is almost equal for beef and poultry. These Marshallian own-price elasiticities are quite different from the static model elasticity estimates.

Table 3.5. Estimated Parameters of Static LAIDS for the Meat Demand in U.S., 1975(1)-2002(4).

Variable	Beef	Pork	Poultry
ln p_1	0.066 (4.91)		
ln p_2	-0.002 (-0.29)	0.048 (5.74)	
ln p_3	0.048 (5.74)	-0.046 (-0.39)	0.002 (8.56)
ln (m/P)	0.089 (1.77)	-0.009 (-0.30)	-0.080 (-2.19)
q_1	0.026 (6.22)	-0.007 (-2.77)	-0.019 (-6.47)
q_2	0.033 (9.04)	-0.021 (9.10)	-0.012 (-4.87)
q_3	0.030 (9.03)	-0.019 (9.40)	-0.011 (-4.66)
T	-0.001 (-23.39)	0.000 (5.33)	0.001 (4.33)
Constant	-0.181 (0.92)	0.286 (2.23)	0.995 (-3.81)

Notes: Homogeneity and symmetry constraints imposed. Poultry estimates derived using adding up constraints. The *t*-values are given in parentheses.

The compensated cross-price elasticities are positive for beef, pork, and poultry, indicating that these meats are all substitutes, presented in upper half of Table 3.8. In particular, a one percent increase in pork price causes a 0.28% increase in beef consumption and a one percent increase in poultry price increases beef consumption by 0.012%. Compensated elasticities from the ECM-LAIDS differ in magnitude, but are similar in that a one percent increase in the price of pork causes a 0.30 % increase in consumption of beef (see the lower half of Table 3.8).

The expenditure elasticity estimates for the static model are presented in the upper half of Table 3.7. The elasticity estimate for beef (1.168) categorizes beef as a “luxury”. The calculated expenditure elasticity estimates are based on the estimates from the static LAIDS model at 0.965 for pork and 0.557 for poultry. This implies that beef is most sensitive to changes in total expenditures, followed by pork, and then poultry. This finding means beef is the biggest gainer (loser) of the three competing meats when consumers increase (decrease) per capita expenditures. Expenditure elasticity estimates from ECM-LAIDS model are reported in lower half of Table 3.7. The estimates 0.66, 1.48, and 1.21 for beef, pork, and poultry respectively, suggest that pork and poultry are "luxury goods" that is when consumers expenditures increase, pork and poultry gain a much higher expenditure share relative to beef.

Table 3.6. Estimated Parameters of ECM-LAIDS for the Meat Demand in U.S., 1975(1)-2002(4).

Variable	Beef	Pork	Poultry
ΔW_{t-1}	-0.097 (1.23)	0.034 (0.50)	0.063 (0.50)
ΔW_{t-2}	-0.269 (3.36)	-0.037 (0.60)	0.307 (2.46)
ΔW_{t-3}	-0.258 (3.45)	-0.131 (2.08)	0.389 (3.25)
ΔW_{t-4}	0.358 (5.17)	0.445 (7.95)	-0.803 (-7.24)
$\Delta \ln p_1$	0.032 (1.10)		
$\Delta \ln p_2$	0.019 (0.55)	0.040 (2.40)	
$\Delta \ln p_3$	-0.043 (-2.47)	-0.050 (-4.41)	0.093 (2.41)
$\Delta \ln (m/P)$	-0.178 (-4.01)	0.139 (4.77)	0.039 (1.37)
q_1	-0.001 (-0.40)	0.004 (1.67)	-0.002 (-0.83)
q_2	0.000 (0.07)	-0.006 (3.64)	0.006 (2.69)
q_3	0.001 (0.40)	-0.001 (0.34)	-0.000 (-0.25)
ECM term	-0.133 (3.36)	-0.291 (-4.62)	0.425 (4.71)

Notes: Homogeneity and symmetry constraints imposed. Poultry estimates derived using adding-up constraints. The *t*-values are given in parentheses. Constant and linear time trend are omitted. The AICC criteria were used to determine the number of lags for budget share variables.

Results suggest that there exist considerable differences in the elasticity estimates generated from the static and dynamic model. However, the estimates from the static model are inaccurate as they do not account for nonstationary properties of the data and as a result, suffer from dynamic misspecification. Further, this study also found evidence that habit persistence plays an important role in the U.S. meat consumption decision-making process.

3.6 Summary and Conclusions

The objective of this paper was to estimate a U.S. meat demand system, using time series techniques. Quarterly per-capita meat disappearance data spanning from 1975(1) to 2002(4) and average retail prices were used. We investigated the time series properties of the data in regard to stationarity and cointegration, and formulated an ECM-LAIDS model consistent with the properties of the data. To facilitate the comparison, both static and ECM-LAIDS models were estimated using seemingly unrelated regression techniques. The study also tested the theoretical restrictions; using likelihood ratio tests, the static model satisfies all the theoretical restrictions of homogeneity and symmetry, but suffers from dynamic misspecification.

Table 3.7. Marshallian and Expenditure Elasticities of the Meat Demand in U.S., 1975(1)-2002(4).

Product	Beef price	Pork price	Poultry price	Expenditure
		Static LAIDS		
Beef	-0.905	-0.059	-0.203	1.168
Pork	-0.001	-0.818	-0.144	0.965
Poultry	-0.266	-0.111	-0.191	0.569
		ECM-LAIDS		
Beef	-0.530	-0.063	-0.063	0.662
Pork	-0.552	-0.739	-0.194	1.486
Poultry	-0.486	-0.222	-0.503	1.212

Note: Derived from homogeneity and symmetry imposed estimates.

Table 3.8. Compensated Price Elasticities of the Meat Demand in U.S., 1975(1)-2002(4).

Product	Beef price	Pork price	Poultry price
		Static LAIDS	
Beef	-0.287	0.275	0.012
Pork	0.508	-0.542	0.033
Poultry	0.034	0.051	-0.086
		ECM-LAIDS	
Beef	-0.180	0.126	0.053
Pork	0.233	-0.314	0.081
Poultry	0.154	0.124	-0.279

Note: Derived from homogeneity and symmetry imposed estimates.

The elasticity estimates obtained were within the ranges reported in a recently published report of U.S. Environmental Protection Agency (EPA), and are consistent with earlier findings of Piggot and Marsh, (2004). The price elasticities estimated from both static and dynamic LAIDS models are close, following the estimates reported in earlier studies by Gao and Shonkwiler (1993), who used the first differenced Rotterdam model to estimate the demand system; Kesavan et al., (1993), who used autoregressive LAIDS and static LAIDS model to estimate the meat demand system. As for the expenditure elasticities, there was a notable difference between the two models. The estimates from the dynamic model are close to those reported in earlier studies by Kesavan et al., (1993) and Wang and Bessler (2003). The evidence of habit persistence found in this study is consistent with earlier findings by Pope et al., (1980), who also found that current period consumption of meat is influenced by previous period meat consumption.

It has been argued in the literature that the estimation of demand systems using econometric techniques that account for unit-roots and cointegration may provide a better framework to incorporate or test theoretical restrictions such as homogeneity and symmetry. This study found that an ECM-LAIDS model for U.S. meats rejected homogeneity. The empirical results in this chapter indicate that the elasticity estimates from the dynamic model differ from the static model but little research is available on estimates reliability. Future research may focus on identifying sources of differing results, perhaps through Monte Carlo or other simulations exercises.

CHAPTER 4

PREDICTIVE ACCURACY OF DYNAMIC U.S. MEAT DEMAND SYSTEMS

4.1 Introduction

In chapter three, the time series properties of aggregate U.S. meat demand data were used to specify a dynamic model (ECM-LAIDS model). This chapter focuses on an empirical evaluation of the ex-post forecasting performance of static and dynamic demand systems for U.S. meat consumption. The aim of this econometric case study is to shed light on the contribution of dynamics to forecasting performance in these demand systems.

The study evaluates the out-of-sample forecasting performance of the static and dynamic LAIDS models, through nominal comparisons of mean square error (MSE) and mean absolute error (MAE) criteria. It also uses MSE-tests to compare forecast reliability. In demand analysis, the conventional MSE and MAE criteria have been used to evaluate the forecasting performance of alternate demand models (Chambers, 1990; Kastens & Brester, 1996). In a recent study, the Ashley, Granger and Schmalensee test (1980) has been used for forecast evaluation (Fraser and Moosa, 2002). More recently, Wang and Bessler (2003) and Robledo and Zapata (2002) adopted the Diebold and Mariano (1995) test to measure predictive accuracy. Despite a large amount of literature on specification and estimation of demand systems, there exists scant focus on forecast MSE-testing. This research provides initial empirical evidence on the forecast performance of the static and dynamic LAIDS models for the U.S. meat demand.

This chapter is organized as follows. Section two describes the empirical models to be estimated. Data and estimation methodology are discussed in section three. Results of out-of-sample forecasts and residual analysis are presented in section four, and finally, summary and conclusions are presented.

4.2 Empirical Model

In chapter three we observed that once the cointegration relationship between the dependent variables and the linear combination of independent variables in the static LAIDS is confirmed, an ECM of the LAIDS can be estimated. The ECM of the LAIDS used in this study follows Karagiannis and Mergos (2002), and as we see in chapter three is given by

$$\Delta w_i = \delta_i \Delta w_{t-1} + \sum_{j=1}^{n} \gamma_{ij} \Delta \ln p_j + \beta_i \Delta \ln(M / P) + \lambda_i \mu_{it-1} + \mu_t . \qquad (1)$$

4.3 Estimation Procedure and Results

Data used in the analysis are quarterly observations over the period 1975(1)-2002(4), providing a total of 112 observations. The quantity data are per capita disappearance data from the United States Department of Agriculture (USDA), the Economic Research Service (ERS) supply, and utilization tables for beef, pork, and poultry (sum of broiler, other-chicken, and turkey). The retail prices used in the study are deflated by nonfood consumer price index (base year being 1984) to obtain the real prices. For a detailed description of the variables and estimation methodology used in this study, refer to chapter three.

4.3.1 Residual Analysis

The residuals from both the static and ECM-LAIDS models were subjected to autocorrelation tests. The residuals from static model were subjected Durbin-Watson test. A test statistic value equivalent or close to two indicates no autocorrelation exists. The results for Durbin-Watson test are reported in Table 4.1. Result from the test confirmed that the static LAIDS model (p-values > 0.05 for beef and pork equations) suffers from autocorrelation suggesting dynamic misspecification. The dynamic model includes lagged budget shares terms as a result the Durbin-Watson test is not applicable. Hence, a white noise test suggested by

Ljung-Box (1978) is used for the autocorrelation test on the residuals of the ECM-LAIDS model the p-values (<0.000) suggest that we reject the null hypothesis of white noise. The AIC criteria were used to identify the best model. However, individual residual series still show a degree of autocorrelation.

Table 4.1. Residual Autocorrelation Test Results for Alternate Specifications of LAIDS Model.

Model	Test Statistic	p-values
Static	Durbin-Watson Statistic	
Beef	0.489	0.545
Pork	0.765	0.243
Dynamic	Ljung-Box portmanteau	
Beef	640.3	<0.000
Pork	431.1	<0.000

4.4 Forecasting Experiment

The difficulty in forecasting with AIDS and other traditional demand systems is that contemporaneous values of price and expenditure are needed to forecast quantities or budget shares. The values of these variables are unknown with the exception of seasonal dummies and constants. In Chambers (1990) and Kastens and Brester (1996), the actual values at $t+1$ were used, which is essentially an out-of-sample fit instead of an out-of-sample forecast. We adopted the same approach, using actual prices and expenditure to generate the forecasts. Once the parameters of each demand model were estimated, we used prices, income, and previous period quantities as predetermined to derive the conditional predictions of the left hand side variables at time $t+1$. In the case of ECM-LAIDS for each commodity, a quantity prediction at the out-of-sample time $t+1$ is given by

$$q_{t+1} = \frac{predLHS + w_{t-1}}{p} x . \qquad (2)$$

The quantity forecasts then were used to generate the forecasting errors.

Diebold and Mariano (1995) suggest that given an actual series and two competing predictions, one may apply a loss criterion (such as MSE or MAE) and then calculate a number of measures of predictive accuracy that allow the null hypothesis of equality in predictive accuracy of the two competing models to be tested. They propose a test statistic that tests null hypothesis of the mean difference between the loss criteria for the two predictions is zero.

The goal is to determine which model provides improved forecasts. Let $\{\hat{e}_{1t}\}_{t=1}^{T}$ and $\{\hat{e}_{2t}\}_{t=1}^{T}$ denote the forecast errors. The subscripts 1 and 2 denote two alternative forecasting equations used to obtain $\{\hat{e}_{1t}\}_{t=1}^{T}$ and $\{\hat{e}_{2t}\}_{t=1}^{T}$, respectively. If $\{e_{1t}^{2}\}$ and $\{e_{2t}^{2}\}$ are the square of $\{\hat{e}_{1t}\}_{t=1}^{T}$ and $\{\hat{e}_{2t}\}_{t=1}^{T}$, the following hypothesis is tested:

$$Ho\text{: } \mathrm{E}\{e_{1t}^{2}\} = \mathrm{E}\{e_{2t}^{2}\} \text{ vs. } \mathrm{E}\{e_{1t}^{2}\}) \neq \mathrm{E}\{e_{2t}^{2}\} . \qquad (3)$$

Diebold-Mariano (1995) reformulated the hypothesis as:

$$Ho_{DM}\text{: } \mathrm{E}\{e_{1t}^{2}\} - \mathrm{E}\{e_{2t}^{2}\} = 0 \text{ vs. } Ha_{DM}\text{: } \mathrm{E}\{e_{1t}^{2}\} - \mathrm{E}\{e_{2t}^{2}\} \neq 0. \qquad (4)$$

To test the null, Diebold-Mariano proposed the test statistic

$$S = [\hat{V}(\overline{d})]^{-1/2}\overline{d} \ , \qquad (5)$$

where, $\overline{d} = \frac{1}{T}\sum_{t=1}^{T}[e_{1t} - e_{2t}]$, $\hat{V}(\overline{d}) = \frac{1}{n}\left[\hat{\gamma}_{0} + 2\sum_{k=1}^{h-1}\hat{\gamma}_{k}\right]$, and $\hat{\gamma}_k = \frac{1}{n}\sum(d_t - \bar{d})(d_{t-k} - \bar{d})$.

We used two approaches to generate out-of-sample forecasts. The first approach used the observations for 1975(1)-1999(4) to estimate the static and dynamic LAIDS model, and subsequently used the estimates to forecast the budget shares over the period 2000(1)-2002(4). The approach is commonly referred to as twelve-step-ahead forecasting. The second approach used the recursive method to generate the multi-step-ahead forecasts. Under this method the sample was divided into estimating and forecasting sub-samples. The estimating sub-sample consisted of the first *T-12* observations, and is used to obtain the initial estimates of the model. The forecasting sub-sample consists of the remaining 12 data points and is used to evaluate forecasting accuracy. The parameter estimates of the models are recursively updated as new observations become available in the forecasting sub-sample. More specifically, out-of-sample forecasts are generated by the following procedure:

- Estimate an initial demand systems for meats using data from 1975(1)-1999(4).
- Forecast one, two, three, and four quarters ahead (out-of-sample) from 2000(1)-2000(4).
- Add a new quarterly observation to the estimation period, and re-estimate the demand system. This is done twelve times, that is, 1975(1)-2000(1), 1975(1)-2000(2), …., 1975(1)-2002(4).
- Forecast one, two, three, and four quarters ahead each time the model is updated, which generates twelve forecast observations for one-step-ahead, and eleven, ten and nine forecast observations for the other three forecast horizons.
- Test the differences in MSE between forecasts from static and dynamic demand systems for each of one, two, three, and four steps-ahead forecasts. Note that the test is based on

twelve observations for the one-quarter-ahead forecasts and nine observations for the four-quarter-ahead forecasts.

Forecast errors were generated comparing the out-of-sample forecasts with the actual values and used to produce two measures (MSE and MAE) of forecasting power. The results are reported in upper half of Table 4.2 and 4.3. The lower values (0.217 and 0.152 MSE for beef and pork, respectively) associated with the dynamic ECM-AIDS model showed that the static LAIDS model is outperformed in out-of-sample forecasting. Additionally we implemented the recently developed Diebold-Mariano (1995) test to compare the predictive accuracy of the two models.

Table 4.2. Twelve-Step Ahead Forecasting Performance of the Alternate Specifications of LAIDS Model.

Meat/Model	MSE	MAE
Beef		
Static	0.764	0.026
Dynamic	0.217	0.011
Pork		
Static	0.235	0.010
Dynamic	0.152	0.009
	Diebold Mariano test	
Beef	6.275	8.866
	(<0.000)	(<0.000)
Pork	0.445	0.804
	(0.656)	(0.421)

Note: All MSE values to be multiplied by10^{-3}. The p-values are enclosed in the parentheses.

The results of the Diebold-Mariano test for the beef and pork equations are presented in the lower half of Table 4.2 and 4.3. The p-values (<0.000) for the Diebold-Mariano test confirm that the dynamic specification outperforms the static model for the beef equation at all forecasting horizons. In the case of pork equation the p-value (0.013) for two-step-ahead forecasting horizon suggests that static model is superior to dynamic model. However, the same cannot be said about three-step-ahead, four-step-ahead and twelve-step-ahead forecasting horizons as the tests (p-values >0.05) fail to reject equality while the MSE (lower values of 0.091, 0.010, and 0.009, respectively) criteria indicates the dynamic model is superior.

Table 4.3. Recursive Forecasting Performance of the Alternate Specifications of LAIDS Model.

Meat/Model	One-step	Two-step	Three-step	Four-step
		MSE		
Beef				
Static	0.547	0.575	0.590	0.585
Dynamic	0.208	0.293	0.112	0.205
Pork				
Static	0.106	0.110	0.109	0.112
Dynamic	0.124	0.207	0.091	0.010
		Diebold-Mariano Test		
Beef	4.443	4.803	6.369	3.715
	(<0.000)	(<0.000)	(<0.000)	(<0.000)
Pork	-1.240	-2.483	1.256	0.631
	(0.214)	(0.013)	(0.209)	(0.527)

Note: All MSE values to be multiplied by10^{-3}. The p-values are enclosed in parentheses.

In addition we evaluated out-of-sample forecast performance using the more recently published data (2003(1)-2004(4)). Following earlier procedures we used the data from 1975(1)-2002(4) as the estimation sample to generate out-of-sample forecasts. The results for the post sample period forecasts are presented in Table 4.4. MSE criteria for beef and pork equation (lower values of 0.501 and 0.348, respectively) suggest that the dynamic model outperforms the static model in forecasting. Further results from the Diebold-Mariano test (p-values < 0.05) confirm that the differences in the MSE criteria are statistically significant for both beef and pork equations suggesting that the dynamic model is superior in forecasting.

Table 4.4. Post Sample Period Forecasting Performance of the Alternate Specifications of LAIDS Model.

Meat/Model	MSE
Beef	
Static	2.848
Dynamic	0.501
Pork	
Static	0.699
Dynamic	0.348
Diebold-Mariano test	
Beef	4.078
	(<0.000)
Pork	2.701
	(0.0069)

Note: All MSE values to be multiplied by10^{-3}. The p-values are enclosed in the parentheses.

4.5 Summary and Conclusions

In this chapter, we examined forecast accuracy of dynamic and static LAIDS models, using U.S. aggregate meat consumption data by means of two approaches. The first approach used the entire forecasting sample to generate forecasts in a single step while the second approach recursively generated one-step-ahead, two-step-ahead, three-step-ahead and four-step-ahead out-of-sample forecasts.

The results from residual analysis indicate that both static and dynamic LAIDS models suffer from autocorrelation problems. However, ECM-LAIDS was shown to be superior in out-of-sample forecasting performance under all the forecasting horizons, as well as the post-sample period forecast evaluation for beef equation. In the case of the pork equation, MSE criteria revealed that the static model was superior in shorter forecasting horizons (one-step-ahead and two-step-ahead) while the dynamic model was superior in the longer forecasting horizons. However, the Diebold-Mariano (1995) test indicated that the difference in MSE values was statistical significant for only the two-step-ahead forecast horizon. The results from post-sample period forecasts suggest that for pork equation, the dynamic model is superior under the conventional MSE criteria, as well as the statistical significance test suggested by Diebold-Mariano.

The findings in this study are consistent with earlier studies by Chamber and Nowman (1997), who used a different data set and found that a dynamic LAIDS model performed better in forecasting when compared to the static LAIDS model; Kastens and Brester (1996), who found that the first differenced LAIDS model was superior in forecasting than an LAIDS model expressed in level; and Wang and Bessler (2003), who found that a dynamic vector error correction model performed better than the static LAIDS model.

The chapter demonstrates that out-of-sample forecasting can be used for model selection. It is easy to implement and can be subjected to formal statistical tests. Although the evidence provided here tends to favor the dynamic LAIDS model, future research may focus on comparing the forecast accuracy for different functional forms using the techniques discussed in the chapter. An example would be an expansion of this research to estimate dynamic demand systems using semiparametric procedures.

CHAPTER 5

SEMIPARAMETRIC CENSORED DEMAND SYSTEMS FOR U.S. MEATS

5.1 Introduction

Food consumption patterns of U.S. households have changed in the last decade, along with the changing socioeconomic composition of the population. Senauer et al. (1991) noted that tastes and preferences for food products are rooted in the fundamental forces of demographics and lifestyles. Households with children present are more likely to prepare food at home than households without children (Kinsey, 1990). Nayga (1995) found that beef and pork expenditures are positively related to household size. He also found that age is significantly related to expenditures on various meat products and expenditure on beef initially increases with age, and then declines. Thus, understanding how demographics shape food consumption is essential in understanding how food policy impacts specific consumer groups. The use of aggregate time series data in chapter three and four to estimate the demand systems does not allow for the measuring of demographic effects. Cross-sectional data collected at the household level provides an opportunity to assess the effects of demographic characteristics.

Empirical analysis of food demand systems increasingly makes use of household-level data that include effects of demographic variables. In parametric specifications, the functional shape of the relationship is predetermined. The quality of the resulting estimator depends heavily on the correctness of this specification. The assumption of a fixed parametric functional form embedded in parametric methods is relaxed in nonparametric models; consequently, there are no parameters to estimate. However, a parametric structure can be combined with nonparametric flexibility to arrive at a more economic theory-consistent specification known as a semiparametric model. In the context of demand analysis, nonparametric smoothing methods have been applied using a single equation framework in studies by Bierens and Pott-Butler (1990), Banks et al., (1997), and Blundell et al., (1998). This econometric analysis compares parametric and semiparametric models for U.S. meat demand.

One reason semiparametric procedures are of research interest lies in the periodical observation that cross-sectional consumption data tend to behave nonlinearly when plotted against demographic variables. In the estimation of LAIDS models, these effects permeate through the system via parameter estimates of the relationship between budget shares, prices and expenditures. The semiparametric LAIDS model can be specified as a parametric LAIDS model with nonparametric effects on demographic variables. Semiparametric procedures are applied to the cross-sectional data since there are a large number of studies on the estimation of demand systems using the cross-sectional data, thereby providing a base for comparison.

The rest of the chapter is organized as follows. Section two presents an econometric model where parametric and semiparametric models are presented. Data source and descriptive statistics are presented in section three. Section four presents a detailed description of estimation methodology. Results obtained from the econometric analysis are summarized in section five. Specifically, a comparison between the parameter estimates and the elasticity estimates from the parametric and semiparametric approaches is provided. The final section includes conclusions and future research implications.

5.2 Econometric model

5.2.1 Parametric Model

In household level survey data, zero-consumption values (non-purchases) are often recorded. Households may have zero-consumption in a given period for various reasons, including sufficient household inventory, nonpreference, and lack of storage. If only observed, positive purchases are used in estimation of demand systems, and OLS estimates will be biased and inconsistent due to sample selectivity bias, thus reducing the predictive power of the model.

Tobin (1958) stated that when estimating relations, where observations in the dependent variable contain zero values, ordinary least-squares estimator produces inconsistent parameter estimates. An explanatory variable in such relationships may be expected to influence both the probability of limit responses and the size of no-limit responses. The Tobit model is specified as follows:

$$q_t = 0 \text{ if } \beta x_t + \xi_t \leq 0, \tag{1}$$

$$q_t = \beta x_t + \xi_t \text{ if } \beta x_t + \xi_t \ \phi \ 0 \ \text{t} = 1,2,\ldots n, \tag{2}$$

where q_t is the dependent variable, x_t is the vector of independent variables and β is the vector of parameters to be estimated, ξ_t is the error term assumed to be normally independent. As can be seen in equations 1 and 2, the decision to consume and the quantity to consume are based on the one set of estimated Tobit coefficients (Haines, Guilkey, & Popkin, 1988). According to Byrne, Capps and Saha (1996), the use of the Tobit model restricts the directional effects to be the same for both participation decision and the expenditure level decision. Thus, the real behavioral pattern is not followed using the Tobit model; the results are not consistent. Therefore, a model that describes the two-step process is needed.

Heckman (1979) proposed a method for dealing with the issue of zero expenditures, modeling the purchase decision using a probit model that determines the probability of purchasing a product. The predictions from the probit model are used to generate the Inverse Mill’s Ratio (IMR), which is the ratio of the estimated values of the standard normal density function to the estimated value of the standard normal cumulative distribution function. The IMR is calculated for each observation in the dataset; mathematically, the Heckman procedures can be described as follows:

$$p^* = x\beta + \varepsilon, \tag{3}$$

$$q^* = x\beta + \mu, \tag{4}$$

$$IMR_{ih} = \frac{\phi(x\beta)}{\Phi(x\beta)}, \tag{5}$$

where equation 3 models the realization of the latent variable p^*, the binary realization variable p takes the values of 1 when $p^*>0$ and takes 0 when $p^* \leq 0$, x is the set of independent variables. In equation 4, q^* contains the information for individuals for which the binary realization equals 1. Following the calculation of IMR, the final equation estimated is augmented with the IMR for correcting the selectivity bias in the demand equation of interest, as described by the following equation:

$$w = f(x\beta) + \lambda IMR, \tag{6}$$

where $f(x\beta)$ is the equation of interest and IMR is the instrumental variable. In the final estimation, only observations with non-limit responses are used. The IMR becomes a variable that links the participation decision with the equation that represents the quantity demanded. According to Heckman (1979), the presence of selectivity bias is found when the parameter λ is statistically significant.

Heien and Wessells (1990) generalized the Heckman two-step procedure to incorporate the IMR for observations with zero values in the dependent variables, thereby using all the observations in the second step. The IMR is calculated for each household (*h)* and commodity *i,* using the maximum likelihood estimates from the probit regression and is therefore the ratio of the standard normal density (ϕ) to the standard normal cumulative density function (Φ):

$$IMR_{ih} = \frac{\phi(x\beta)}{\Phi(x\beta)} \text{ for } y_{ih} = 1, \tag{7}$$

and

$$IMR_{ih} = \frac{\phi(x\beta)}{1-\Phi(x\beta)} \text{ for } y_{ih} = 0, \tag{8}$$

where x consists of explanatory variables including demographics, and log transformed prices, and y_{ih} is a dummy variable where $y_{ih} = 1$ if the household consumes the i^{th} good and $y_{ih} = 0$ if the household does not consume the i^{th} good. We use the generalized version, as it allows inclusion of all the observations in the second step regression.

In the second stage, we use the AIDS model from Deaton and Muellbauer (1980). The general specification of the AIDS model is given by

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(M / P) \tag{9}$$

where w_i is the share associated with the i^{th} good, α_i is the constant coefficient in the i^{th} share equation, γ_{ij} is the slope coefficient associated with the j^{th} good in the i^{th} share equation, p_j is the price of the j^{th} good, M is the total expenditure on the system of goods given by the following equation: $M = \sum_{i=1}^{n} p_i q_i$ where q_i is the quantity demanded for the ith good. P is a general price index defined by:

$$\log P = \alpha_0 + \sum_{i=1}^{n} \alpha_i \log p_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \log p_i \log p_j \,. \tag{10}$$

To comply with the theoretical properties of consumer theory, the following restrictions are imposed on the parameters in the AIDS model:

- *Adding up restriction*: $\sum_{i=1}^{n} \alpha_i = 1, \sum_{i=1}^{n} \beta_i = 0, \sum_{i=1}^{n} \gamma_{ij} = 0$, allowing the budget share to sum to unity;
- *Homogeneity*: $\sum_{j=1}^{n} \gamma_{ij} = 0$, which is based on the assumption that a proportional change in all prices and expenditure does not affect the quantities purchased. In other words, the consumer does not exhibit money illusion;
- *Symmetry*: $\gamma_{ij} = \gamma_{ji}$, represents consistency of consumer choices.

In empirical studies, to avoid non-linearity and reduce the multicolinearity effects in the model, equation (2) is sometimes approximated by a Stone index defined as $\log P = \sum_{i=1}^{n} w_i \log p_i$. We use the simple linear AIDS model in our empirical investigation and augment the model to incorporate the demographic variables (d_k), and the selectivity bias correction term (IMR) is given by

$$w_i = \alpha_i + \sum_{k}^{s} \rho_{ik} d_k + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(M / P) + \lambda_i IMR_i . \tag{11}$$

The flexible functional form of the LAIDS model allows us to easily carry out the elasticity analysis. The demand elasticities are calculated as functions of the estimated parameters, and they have standard implications. According to Green and Alston (1990), elasticities in LAIDS can be expressed as: $\eta_i = 1 + \beta_i / w_i$ for income elasticity and $\eta_{ij}^{*} = -\delta_{ij} + w_j + \gamma_{ij} / w_i$ for compensated elasticity. The uncompensated elasticities are computed from the formula $\eta_{ii} = -\delta_{ij} - \beta_i + \gamma_{ij} / w_i$, and $\eta_{ij} = \gamma_{ij} / w_i - \beta_i \ w_j / w_i$.

5.2.2 Semiparametric Model

The parametric LAIDS model ignores the periodical observation that cross-sectional consumption data patterns, particularly in relation to demographics, tend to behave nonlinearly. Hence, the parameter estimates obtained through parametric models may not be the most accurate representation of the relationship between budget shares, prices, and expenditures. To overcome these problems, the semiparametric LAIDS model is specified with parametric LAIDS effects and nonparametric effects between budget shares and demographics. In a semiparametric specification, the functional form is relaxed by specifying parametric and nonparametric effects. The strength of this method lies in the fact that one does not need to specify a parametric form for the nonlinearity part and is computationally much easier than most of nonlinear regression models (Bachmeier & Li, 2002). The semiparametric LAIDS model is given by

$$w_i = \alpha_i + \sum_{k}^{s} \mu(d_k) + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(M / P) + \lambda_i IMR_i , \tag{12}$$

where d_k is a demographic variable. The theoretical restrictions of adding-up, homogeneity, and symmetry conditions can be imposed on the semiparametric model also.

The first step semiparametric probit model is specified using log prices as a parametric component, while the demographic variables (age of reference person and household size) are included as additive nonparametric components. The econometric representation of the semiparametric probit model is given by

$$q_i^* = \alpha_i + \beta_1 \ln pb + \beta_2 \ln pp + \beta_3 \ln pc + \beta_4 \ln ps + f(age) + f(hsize) + \varepsilon_i , \tag{13}$$

$$q_i^* = \alpha_i + \beta_1 \ln pb + \beta_2 \ln pp + \beta_3 \ln pc + \beta_4 \ln ps + f(age, hsize) + \varepsilon_i . \tag{14}$$

The predicted values obtained from the probit model are used to construct the IMR. The semiparametric LAIDS model (Eq. 15) can be estimated using the *double residual* approach outlined by Robinson (1988). Semiparametric seemingly unrelated regression techniques are used to estimate the demand system. Based on the model described by Robinson (1998), the parametric LAIDS model can be expressed as a semiparametric LAIDS model as follows:

$$w_i = \alpha_i + \gamma_1 \ln pb + \gamma_2 \ln pp + \gamma_3 \ln pc + \gamma_4 \ln ps + \gamma_5 \ln xp \\ + f(age) + f(hsize) + \lambda_i IMR + \mu_i, \tag{15}$$

$$w_i = \alpha_i + \gamma_1 \ln pb + \gamma_2 \ln pp + \gamma_3 \ln pc + \gamma_4 \ln ps + \gamma_5 \ln xp \\ + f(age, hsize) + \lambda_i IMR + \mu_i. \tag{16}$$

Table 5.1 Definitions of the Variables Included in the Model.

Variable	Description
ln pb	Log transformed beef price
ln pp	Log transformed pork price
ln pc	Log transformed poultry price
ln ps	Log transformed seafood price
ln xp	Log of expenditure
Age	Age of reference person
Hsize	Household size

Taking the conditional expectations for both sides of equation (15) will result in the equation below:

$$E(w_i \mid age, hsize) = \gamma_1 E(\ln pb \mid age, hsize) + \gamma_2 E(\ln pp \mid age, hsize) + \gamma_3 E(\ln pc \mid age, hsize) \\ + \gamma_4 E(\ln ps \mid age, hsize) + \gamma_5 E(\ln xp \mid age, hsize) + f(age) + f(hsize). \tag{17}$$

Subtracting equation (17) from equation (15) results in equation (18) as follows:

$$w_i - E(w_i \mid age, hsize) = \gamma_1 (\ln pb - E(\ln pb \mid age, hsize)) + \gamma_2 (\ln pp - E(\ln pp \mid age, hsize)) \\ + \gamma_3 (\ln pc - E(\ln pc \mid age, hsize)) + \gamma_4 (\ln ps - E(\ln ps \mid age, hsize)) \\ + \gamma_5 (\ln xp - E(\ln xp \mid age, hsize)) + \lambda_i IMR. \tag{18}$$

By assuming the right hand side variables to be y_i (representing the budget shares after removing the nonparametric effects) and left hand side variables to be x_i (representing the price and expenditure variables obtained after removing the nonparametric effects) equation (18) can be written as a system of equations and estimated using the seemingly unrelated regression procedure. The system of equations using matrix notation is given by,

$$\begin{bmatrix} y_b \\ y_p \\ y_c \\ y_s \end{bmatrix} = \begin{bmatrix} x_b & 0 & 0 & 0 \\ 0 & x_p & 0 & 0 \\ 0 & 0 & x_c & 0 \\ 0 & 0 & 0 & x_s \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_5 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}. \qquad (19)$$

We can further simplify Eq. 19 using more compact notation as

$$Y = X\gamma + \varepsilon, \qquad (20)$$

by obvious substitution. In terms of the equation errors ε_1, ε_2, ε_3, and ε_4, the assumptions employed are as follows:

1. All errors have zero mean: E(ε_i) = 0; for i = 1,2,3,4;
2. In a given equation the error variance is constant over time, but each equation can have a different variance;
3. Two errors in different equations but corresponding to the same household are correlated;
4. Errors in different households are not correlated.

Univariate and bivariate loess smoothing techniques are used for the estimating the probit model and the semiparametric LAIDS model. Loess, originally proposed by Cleveland (1979) and further developed by Cleveland and Devlin (1988), specifically denotes a method that is (somewhat) more descriptively known as locally weighted polynomial regression. At each point in the data set, a low-degree polynomial is fit to a subset of the data, with explanatory variable values near the point where response is being estimated. The polynomial is fit using weighted least-squares, giving more weight to points near the point whose response is being estimated and less weight to points further away. The value of the regression function for the point is then obtained by evaluating the local polynomial using the explanatory variable values for that data point. The loess fit is complete after regression function values have been computed for each of the n data points.

5.3 Data

The data used in this study are extracted from the diary component of the 2003 U.S. Consumer Expenditure Survey (CES) of the Bureau of Labor Statistics. The CES was designed to collect information on expenditures incurred by respondents during a survey week. The respondents are part of a national probability sample of households designed to represent the total civilian population. The total number of households (diaries) in the survey was 15,827. However, households which reported incomplete socioeconomic and demographic information were dropped from the analysis. In addition, households with negative incomes were deleted from the sample. Consequently, the number of households analyzed in the study is 5,454. In this study, we consider expenditures for four meat categories: beef, pork, poultry, and seafood. Budget shares for each category were calculated, based on the weekly expenditures reported by the households. Preliminary analysis of the data involved generating descriptive statistics for the sample data. The descriptive statistics of the budget shares reported in Table 5.2 indicate that the mean of household expenditure share on beef is the highest at 34%, followed by pork (25%), poultry (23%) and seafood (17%). The variance measure indicates that the beef expenditure share exhibits the highest variability, followed by poultry, pork and seafood expenditure shares. In the case of non-zero observation, the mean expenditure share for beef (52%) was again the highest, yet the mean expenditure shares for rest of the meats were very close to one another. The standard deviation measure for non-zero observations did not vary much. A large numbers of households with zero expenditure on each of the meat items were reported. Hence, there existed a need to use specialized censoring techniques in the estimation of the demand system.

Table 5.2 Descriptive Statistics of Budget Shares for Beef, Pork, Poultry, and Seafood in the U.S., 2003.

Variable	Mean	Std.Dev	Number	Mean	Std.Dev
	All Observations			Non-Zero Observations only	
Beef share	0.341	0.335	3543	0.523	0.277
Pork share	0.252	0.297	3186	0.432	0.272
Poultry share	0.237	0.304	2974	0.435	0.289
Seafood share	0.169	0.274	2234	0.413	0.288
Total	5,454				

The main drawback in using the CES data is the lack of price or quantity information concerning expenditure items. However, the meat expenditure categories and regional assignments of the CES match those assigned to the consumer price index (CPI) for each aggregate meat category. Therefore, households are matched to the appropriate regional price index using information regarding the household's regional affiliation and month of the interview. Table 5.3 provides descriptive statistics of the CPI and demographic characteristics which are used as explanatory variables in the study. The descriptive statistics for explanatory variables indicate that the mean price for beef, pork, poultry, and seafood were 2.73 $/lb, 3.09 $/lb, 1.046 $/lb and 1.89 $/lb respectively. The mean household size was 2.87 and mean age of the reference person was 48.40. The standard deviation measure indicates poultry price showed the highest variation, followed by beef, poultry, and seafood. The average household size and age were 2.87 and 48.39, respectively.

Table 5.3. Descriptive Statistics of Prices, Total Expenditure, Household size, and Age in the U.S., 2003.

Variable	Mean	Std.Dev	Minimum	Maximum
Beef price ($/lb)	2.723	0.157	2.424	3.034
Pork price ($/lb)	3.096	0.225	2.564	3.514
Poultry price ($/lb)	1.046	0.119	0.855	1.273
Seafood price ($/lb)	1.899	0.018	1.869	1.925
Expenditure ($/week)	22.867	36.058	10.05	203.852
Household size	2.871	1.47	1.0	17.0
Age	48.391	15.67	16.0	85.0

Note: Prices are expressed in log transformation.

5.4 Estimation Methodology

5.4.1 Parametric Method

In the estimation of the parametric model, the first step involved a parametric probit regression for each equation in the system, where zero consumption was observed in at least one observation. The dependent variable in the probit regression was set to one, if the good is consumed in strictly positive quantities and zero otherwise. Independent variables included log transformed prices and demographic variables such as household size and age of the reference person. A linear relationship was assumed between the dependent variables and the two demographics variables. The predicted values from the probit regression were used to construct the IMR based on Eq.7 and Eq.8. The computed IMRs were then used as instruments in the second-stage estimation of the full-demand system.

In the second stage, the augmented LAIDS model (Eq. 11) is estimated using iterated seemingly unrelated regression procedures to gain efficiency and to account for the possible contemporaneous correlation among the error terms. The theoretical restrictions of adding up, symmetry, and homogeneity were imposed as restrictions on the parameters to be estimated. The parameter estimates thus obtained were used to generate the elasticity estimates, using the Green and Alston (1990) formula.

5.4.2 Semiparametric Method

The estimation of a semiparametric model involved the use of a first step partial linear probit model for each equation in the system. The dependent variable remained the same as the above parametric probit regression. The S-Plus statistical package provided a GAM function to fit generalized additive models. Estimating the smoothness of a relationship requires two decisions: the type of smoother and the size of the neighborhood (bandwidth). We used loess, a locally-weighted regression smoother and the bandwidth was selected by cross-validation method built into the procedure. The predicted values from the partial linear probit regression were used to construct the IMR based on Eq.7 and Eq.8. The computed IMRs were then used as instruments in the second-stage estimation of the semiparametric LAIDS model.

In the second stage, the augmented semiparametric LAIDS model is estimated using Robinson (1988) approach. Following Robinson (1988), the steps below are carried out in estimating $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \lambda_i$, *f(age)*, and *f(hsize)*:

1. The unknown conditional means, in equation (17) for each share equation are estimated, using nonparametric estimation techniques;
2. These estimates are substituted in place of the unknown functions in equation (18), and the coefficients $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \lambda_i$ are estimated, using the seemingly unrelated regression techniques;

Substitute the estimated $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \lambda_i$ into share equations and estimate *f(age)*, and *f(hsize)*.

5.5 Empirical Results

5.5.1 Parametric Model

The results from the parametric approach to the two-step estimation of the demand system are presented in Tables 5.4 and 5.5. Results from the first step probit regression (Table 5.4) indicate that demographic variables for family size and age of the reference person are highly significant, suggesting an importance in meat consumption decisions. The second stage results (Table 5.5) indicate significant parameter estimates for the IMR in each equation, suggesting that if we ignore zero expenditure, there exists a strong selectivity bias. Additionally, a majority of the parameter estimates were found to be significant. The overall fit indicated by adjusted R-square suggests that the model fit is good for all the share equations. Household size is significant and positively related to expenditure shares on beef, pork, and poultry, but is negatively related to seafood. Age of the reference is also significantly related to the expenditure shares of all the meat products. Expenditure shares of beef and poultry were found to be negatively related to the age. However, pork and seafood budget shares were found to be positively related to the age.

5.5.2 Semiparametric Model

In the semiparametric estimation preliminary data analysis was carried out to explore the relationship between each of the budget shares and demographic variables involving age and household size. Nonparametric regression techniques (loess smoother) were used to generate the fitted values for budget shares. Three dimensional surface plots were created using these fitted values. The surface plot (Fig. 5.1) shows a nonlinear relationship exists between beef budget shares and demographic variables of age and household size. The expenditure share of beef exhibits a distinctive, nonlinear pattern as the household size increases, holding age constant.

Similarly, nonlinear patterns, although not that distinctive, exist when age increases with the household size held constant. Figures 5.2 and 5.3 for pork and poultry expenditures also exhibit completely opposite and distinctive nonlinearities. In Fig. 5.4, the nonlinearities for seafood expenditures with respect to age and household size, although not as distinctive as beef expenditure, are similar in shape. Overall, the nonlinearities with respect to age and household size vary across the share equations.

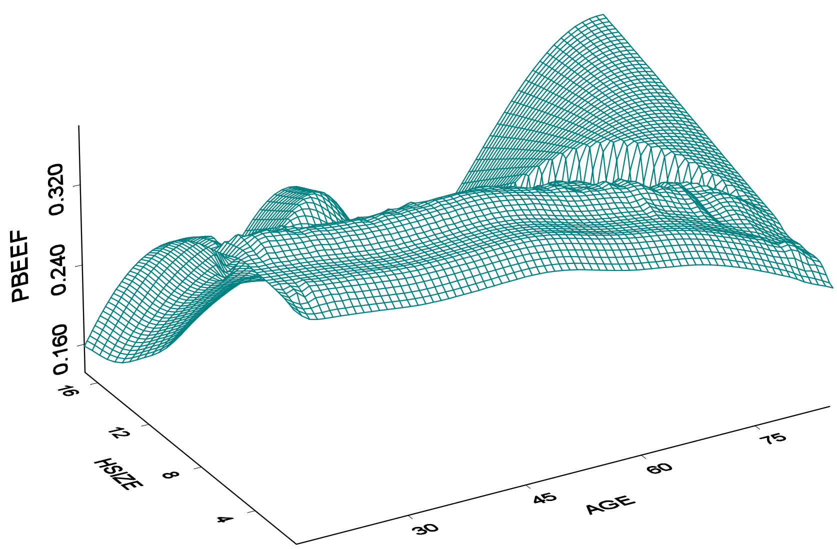

Figure 5.1. Nonparametric Estimates of Age and Household size for Beef Budget Shares.

The results from an additive semiparametric approach to the two-step estimation of the demand system are presented in Tables 5.6 and 5.7. The estimates from the first-step semiparametric probit model (Table 5.6), where age of the reference person and household size are included as nonparametric components, are used to construct the IMR. The parameter estimates from the semiparametric probit model are quite close to their parametric counterparts.

The results from the semiparametric LAIDS model (Table 5.7) are similar to their parametric equivalents in terms of significant parameter estimate for the IMR and significance of parameter estimates. The similarities in results suggest that a nonparametric treatment of demographic variables (age and household size) has little impact on model fit and parameter estimates.

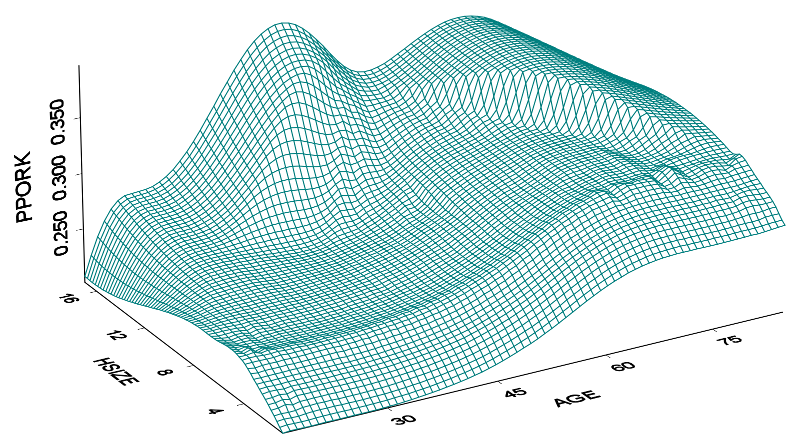

Figure 5.2. Nonparametric Estimates of Age and Household size for Pork Budget Shares.

The results from the bivariate semiparametric approach to the two-step estimation of the demand system are presented in Table 5.8. The results from the semiparametric LAIDS model (Table 5.8) are similar to their parametric and additive semiparametric equivalents, in terms of significant parameter estimates for the IMR and significance of parameter estimates. The similarities in results suggest that a bivariate nonparametric treatment of demographic variables (age and household size) has little impact on the model fit and parameter estimates. The similarities in the results from the additive and bivariate smoothing techniques suggest there is not much to gain from bivariate smoothing of age and household size.

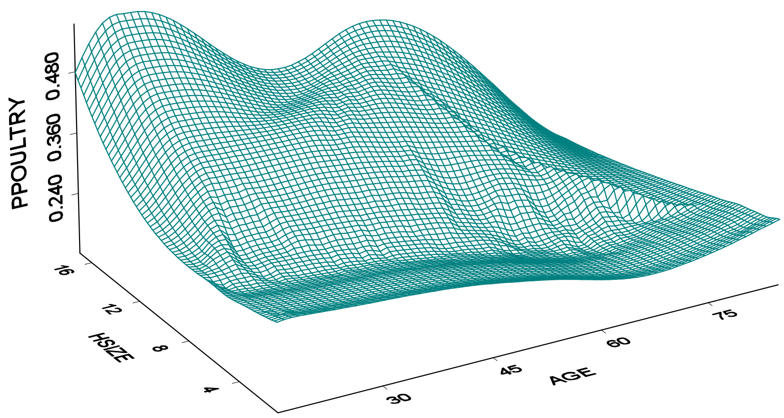

Figure 5.3. Nonparametric Estimates of Age and Household size for Poultry Budget Shares.

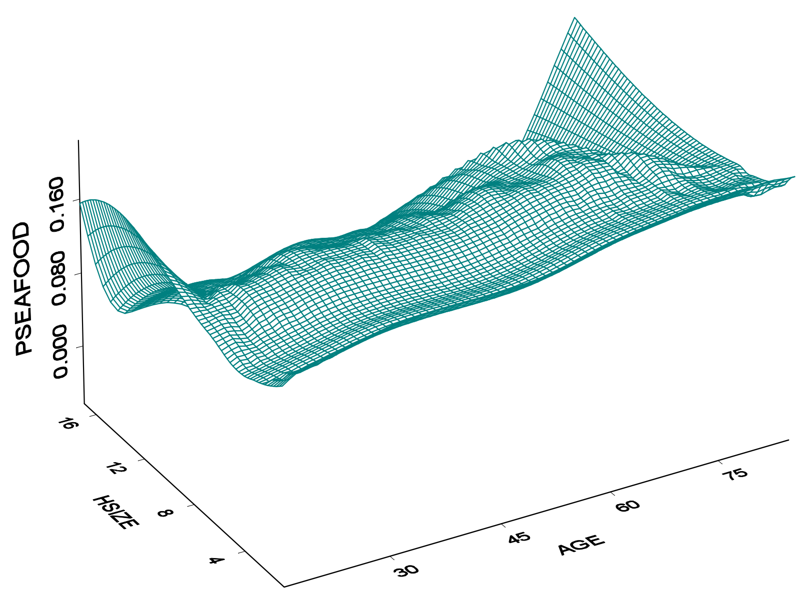

Figure 5.4. Nonparametric Estimates of Age and Household size for Seafood Budget Shares.

5.5.3 Residual Analysis

The residuals obtained from parametric probit models were subjected to further analysis. The histogram density plots of the residuals from the parametric probit model indicate that the assumption of normality is clearly violated raising doubts about the parametric probit model. A clear trend of heavy concentration of residuals around a value of one and below zero values was observed for each of the probit models.

Table 5.4. Estimated Parameters of First Stage Parametric Probit Model for Beef, Pork, Poultry, and Seafood Demand in the U.S., 2003.

Variables	Beef	Pork	Poultry	Seafood
Constant	22.205	-18.007	5.481	28.184
	(9.916)	(9.710)	(9.645)	(9.684)
Beef price	-0.356	0.290	0.396	0.731
	(0.373)	(0.365)	(0.362)	(0.365)
Pork price	0.274	0.455	-0.151	-0.207
	(0.241)	(0.235)	(0.235)	(0.235)
Poultry price	-0.315	-0.605	-0.066	0.191
	(0.181)	(0.178)	(0.176)	(0.177)
Seafood price	-3.819	3.086	-1.221	-6.194
	(1.918)	(1.875)	(1.863)	(1.871)
Age	-0.003	0.005	-0.005	0.003
	(0.001)	(0.001)	(0.001)	(0.001)
Household size	0.104	0.115	0.083	0.045
	(0.013)	(0.012)	(0.012)	(0.012)

Note: The standard errors are enclosed in parentheses.

The residuals obtained from semiparametric probit models were also subjected to a similar analysis. The histogram density plots of the residuals from the semiparametric probit model also indicate that the assumption of normality is clearly suspect. Moderate differences in terms of heavy concentration of residuals around one and below zero were noticed for each of the probit models. The residuals obtained from parametric and semiparametric LAIDS models were also analyzed for distributional assumptions. The density plots of the residuals obtained from the parametric LAIDS show that the normality assumption does not hold. An important feature of the density plots that is common to all the share equations is that they all appear to be skewed to the left, i.e., the left tail is longer.

Table 5.5. Estimated Parameters of Parametric LAIDS Model for Beef, Pork, Poultry, and Seafood Demand in the U.S., 2003.

Variable	Beef	Pork	Poultry	Seafood
Beef price	-0.151			
	(0.051)			
Pork price	0.067	0.027		
	(0.028)	(0.033)		
Poultry price	-0.007	-0.116	0.027	
	(0.022)	(0.018)	(0.023)	
Seafood price	0.090	0.096	-0.208	0.022
	(0.067)	(0.047)	(0.060)	(0.096)
Expenditure	-0.118	-0.118	-0.082	0.312
	(0.002)	(0.003)	(0.003)	(0.009)
Household size	0.018	0.022	0.006	-0.046
	(0.003)	(0.001)	(0.002)	(0.007)
Age	-0.001	0.001	-0.001	0.001
	(0.0001)	(0.0001)	(0.0001)	(0.0006)
IMR	0.342	0.283	0.326	-0.952
	(0.003)	(0.003)	(0.003)	(0.006)
Constant	0.576	0.235	0.532	-0.348
	(0.034)	(0.027)	(0.022)	(0.057)
Adj. R-square	0.638	0.632	0.549	

Notes: Homogeneity and symmetry constraints imposed. Seafood estimates derived using adding up constraints. The standard errors are enclosed in parentheses.

The residuals obtained from the semiparametric LAIDS model were subject to further analysis. The density plots of the residuals obtained from the parametric LAIDS model showed that the normality assumption does not hold. The density plots of residuals from all the share equations were similar to the above plots from the parametric model i.e., appear to be skewed to the left.

Table 5.6. Estimated Parameters of a First Stage Semiparametric Additive Probit Model for Beef, Pork, Poultry, and Seafood Demand in the U.S., 2003.

Variables	Beef	Pork	Poultry	Seafood
Constant	21.642	-18.733	5.752	27.808
	(9.936)	(9.758)	(9.679)	(9.713)
Beef price	-0.322	0.332	0.403	0.755
	(0.373)	(0.367)	(0.363)	(0.366)
Pork price	0.269	0.449	-0.152	-0.209
	(0.241)	(0.236)	(0.236)	(0.237)
Poultry price	-0.293	-0.581	-0.058	0.194
	(0.181)	(0.179)	(0.177)	(0.178)
Seafood price	-3.763	3.159	-1.286	-6.151
	(1.919)	(1.884)	(1.869)	(1.875)

Note: The standard errors are enclosed in parentheses.

5.5.2 Elasticity Analysis

The parameter estimates obtained from both the parametric and semiparametric models were later used to generate the Marshallian and compensated price elasticities and expenditure elasticities reported in Tables 5.8-5.9. The estimated Marshallian own-price elasticities from the parametric model found in upper half of Table 5.9, were all negative for beef, pork, poultry, and seafood (-1.256, -0.661, -0.743, and -1.387, respectively) suggesting downward sloping demand curve. These results meant that seafood and beef consumption were more sensitive to own price changes (price elastic), while pork consumption was least sensitive to changes in its own price

(price inelastic). The expenditure elasticity estimates found in the upper half of Table 5.9, were 0.671 for beef, 0.532 for pork, 0.653 for poultry, and 2.842 for seafood. This implies that seafood was the most sensitive to changes in total expenditures, followed by beef, poultry, and pork. This finding means seafood was the biggest gainer (loser) of the three competing meats when consumers increased (decreased) expenditures on meat.

Table 5.7. Estimated Parameters of the Semiparametric Additive Approach to the LAIDS Model for Beef, Pork, Poultry, and Seafood Demand in the U.S., 2003.

Variable	Beef	Pork	Poultry	Seafood
Beef price	-0.128			
	(0.051)			
Pork price	0.067	0.026		
	(0.028)	(0.034)		
Poultry price	-0.005	-0.111	0.021	
	(0.022)	(0.019)	(0.023)	
Seafood price	0.066	0.017	0.095	-0.179
	(0.067)	(0.059)	(0.047)	(0.059)
Expenditure	-0.112	-0.118	-0.081	0.312
	(0.003)	(0.003)	(0.003)	(0.009)
IMR	0.342	0.282	0.326	-0.915
	(0.003)	(0.003)	(0.004)	(0.006)
Adj. R-square	0.637	0.629	0.547	

Notes: Homogeneity and symmetry constraints imposed. Seafood estimates derived using adding up constraints. The standard errors are given in parentheses.

The Marshallian own-price elasticities generated from the semiparametric additive approach were -1.189, -0.665, -0.772 and -1.214 for beef, pork, poultry, and seafood, respectively, found in the middle of Table 5.9. These results meant that seafood and beef consumption was more sensitive to own price changes, while pork consumption was least sensitive to changes in its own price. The expenditure elasticity estimates calculated based on the estimates from the semiparametric additive LAIDS model were 0.668 for beef, 0.532 for pork, 0.656 for poultry, and 2.844 for seafood (lower half of Table 5.9). This implied that seafood is the most sensitive to changes in total expenditures, followed by beef, poultry, and pork. This finding meant seafood was the biggest gainer (loser) of the three competing meats when consumers increased (decreased) expenditures on meat. These results show that there is not much difference among the elasticity estimates generated from parametric and semiparametric estimation.

Table 5.8. Estimated Parameters of the Semiparametric Bivariate Approach to the LAIDS Model for Beef, Pork, Poultry, and Seafood Demand in the U.S., 2003.

Variable	Beef	Pork	Poultry	Seafood
Beef price	-0.128			
	(0.051)			
Pork price	0.069	0.028		
	(0.028)	(0.033)		
Poultry price	-0.003	-0.111	0.021	
	(0.022)	(0.018)	(0.023)	
Seafood price	0.062	0.013	0.094	-0.169
	(0.067)	(0.059)	(0.047)	(0.059)
Expenditure	-0.112	-0.118	-0.081	0.312
	(0.003)	(0.003)	(0.003)	(0.009)
IMR	0.342	0.282	0.326	-0.951
	(0.003)	(0.003)	(0.004)	(0.006)
Adj. R-square	0.637	0.629	0.548	

Notes: Homogeneity and symmetry constraints imposed. Seafood estimates derived using adding up constraints. The standard errors are given in parentheses.

The compensated cross-price elasticities from the parametric LAIDS model were positive for beef, indicating a substitution relationship with pork, poultry, and seafood (upper half of

Table 5.10). In particular, a one percent increase in pork price caused a 0.53% increase in beef consumption, while a one percent increase in poultry price increased beef consumption by 0.26%. A complementary relationship was found for pork with poultry and seafood consumption, while seafood and poultry were found to have a substitution relationship.

Table 5.9. Marshallian and Expenditure Elasticities of Meat Demand in the U.S., 2003.

Product	Beef price	Pork price	Poultry price	Seafood Price	Expenditure
			Parametric Approach		
Beef	-1.256	0.359	0.110	0.115	0.671
Pork	0.531	-0.661	-0.273	-0.128	0.532
Poultry	0.164	-0.321	-0.743	0.247	0.653
Seafood	-0.506	-0.775	-0.173	-1.387	2.842
			Semiparametric Additive Approach		
Beef	-1.189	0.361	0.117	0.042	0.668
Pork	0.533	-0.665	-0.255	-0.144	0.532
Poultry	0.172	-0.302	-0.772	0.246	0.656
Seafood	-0.655	-0.800	-0.173	-1.214	2.844
			Semiparametric Bivariate Smoothing Approach		
Beef	-1.189	0.368	0.122	0.030	0.668
Pork	0.542	-0.656	-0.255	-0.162	0.532
Poultry	0.178	-0.303	-0.774	0.241	0.657
Seafood	-0.679	-0.827	-0.181	-1.155	2.843

Note: Derived from the homogeneity and symmetry imposed estimates.

The compensated cross-price elasticities generated from the semiparametric LAIDS model were positive for beef, indicating a substitution relationship with pork, poultry, and seafood (lower half of Table 5.10). In particular, a one percent increase in pork price caused a 0.53% increase in beef consumption, while a one percent increase in poultry price increased beef consumption by 0.27%. A complementary relationship was found for pork with poultry and seafood consumption, while seafood and poultry were found to have a substitution relationship. Overall, the signs remained the same, but magnitudes of some cross-price elasticities differed from their parametric counterparts.

Table 5.10. Compensated Price Elasticities of Meat Demand in the U.S., 2003.

Product	Beef price	Pork price	Poultry price	Seafood Price
		Parametric Approach		
Beef	-1.028	0.528	0.269	0.229
Pork	0.712	-0.526	-0.147	-0.038
Poultry	0.386	-0.156	-0.588	0.358
Seafood	0.461	-0.057	0.501	-0.905
		Semiparametric Additive Approach		
Beef	-0.962	0.530	0.276	0.155
Pork	0.714	-0.531	-0.128	-0.054
Poultry	0.395	-0.137	-0.616	0.357
Seafood	0.312	-0.081	0.501	-0.732
		Semiparametric Combined Approach		
Beef	-0.961	0.537	0.280	0.143
Pork	0.724	-0.522	-0.128	-0.072
Poultry	0.402	-0.137	-0.618	0.352
Seafood	0.288	-0.108	0.493	-0.673

Note: Derived from the homogeneity and symmetry imposed estimates.

5.6 Summary and Conclusions

Household food expenditures on four meat categories were analyzed using the cross-sectional data extracted from the 2003 U.S. Consumer Expenditure Survey. Modeling demographic variables, age of the reference person, and household size was of interest. Initial exploratory analysis revealed nonlinear relationships existed between budget shares and demographic variables consistent with earlier finding by Nayga (1995). Hence, a more flexible, semiparametric approach was also used to estimate the demand system.

The two-step estimation procedure proposed by Heckman (1979) was used to estimate the censored demand system. The first step involved estimation of parametric and semiparametric probit regressions to model the consumer purchase decisions. The residuals from both parametric and semiparametric probit model violated the normality assumption and are skewed to the left. This could be due to the large proportion of zero expenditures and single meat product consumption by a household. These results also mean that the conventional probit model is unable to capture the data patterns adequately and that alternative modeling strategies should be explored in future work.

In the second step, parametric and semiparametric specifications of the LAIDS model were used to estimate a demand system. The results from these estimation methods indicated there was not much difference in terms of model fit, but there existed slight differences in magnitudes of elasticity estimates. These results suggest that the two-step approach based on Robinson (1988), adopted to estimate the semiparametric model, does not seem to capture the nonlinearities between budget shares and demographic variables found in the initial exploratory analysis. Elasticity estimates generated in this study are consistent with earlier studies by Huang and Haidacher (1983) and Park and Rodney (1996) who used Nationwide Food Consumption survey data to estimate the demand system.

Some suggestions for future research include the following. Specification of the first-step probit regressions using formal model specification tests relative to including variables in an ad hoc manner. In addition, future research can consider multivariate modeling of the decision to purchase, using parametric and semiparametric techniques. Another topic that can be addressed in future research is the comparison of the LAIDS demand system and other alternative demand models semiparametrically. As multivariate semiparametric testing procedures are developed, model specification tests should be used to discriminate among alternative functional forms.

CHAPTER 6

SUMMARY, CONCLUSIONS AND FUTURE RESEARCH

This dissertation has analyzed results from empirical estimations of Linear Almost Ideal Demand Systems (LAIDS) for U.S. meats. The dissertation was motivated by two research themes. The first research theme was a time series analysis of nonstationarity in demand systems using cointegration theory. The second research theme focused on developing a flexible semiparametric LAIDS model for estimating elasticities from cross-sectional data and permitted demographic effects, such as age of the head of the household and household size, to enter the model nonparametrically. To accomplish this task, the dissertation was structured as follows. The second chapter provided a condense review of literature concerning the estimation of demand systems. Similar to the structure of the dissertation, the review was structured to include empirical work on static and dynamic models using time series data, and on the estimation of demand systems using cross-sectional data. This review identified the AIDS model as one of the most popular demand systems used in applied demand analysis in agricultural economics. AIDS models estimated with time series data were used for two general purposes: a) to estimate price and expenditure elasticities (some studies estimate Marshallian and compensated price elasticities) and b) to forecast quantities consumed and budget shares. Cross-sectional data analyses reviewed in the literature were then used to calculate elasticities. Building upon this background, the dissertation proceeded to develop three econometric case studies on the estimation of LAIDS models for U.S. meats (beef, pork, poultry and seafood).

The first case study reported in Chapter Three provided an analysis of unit-roots and cointegration in the estimation of a LAIDS model. The data included budget shares, prices and expenditures for beef, pork and poultry and were quarterly from 1975(1)-2002(4). Empirical tests of seasonal unit-roots were applied to each series to diagnose the nature of seasonal behavior in the consumption data. It was found that each series contained one-unit root but no seasonal-unit roots were found. Therefore, the LAIDS model was specified as a dynamic error-correction model (ECM) of the Engle-Granger type (e.g., Karagiannis et al, 2002) with a general lag structure that allowed for habit persistence. Empirical work with demand systems has often reported the rejection of demand restrictions such as homogeneity and symmetry, and in some works, it has been argued that nonstationarity and dynamics may help explain such theoretical properties of demand systems. The range of values in elasticity estimates was consistent those reported in previous work (e.g., Gao & Shonkwiler, 1993; Kesavan et al., 1993; Piggot et al, 2004). Habit persistency was also found to be consistent with findings in previous work (e.g., Pope et al., 1980). This case study reported that, although homogeneity was not rejected in a static model of U.S. meats using quarterly time series data, it was rejected in its dynamic counterpart. Therefore, even in the case where dynamic misspecification of a demand system is accounted for, this may not be sufficient to estimate theoretically consistent demand systems.

The case study in Chapter Four was an extension of the first case study. The focus of empirical analyses on demand systems has either been on the estimation of demand elasticities (as in the first case study) or on generating commodity forecasts. This case study focused on the forecasting performance of static and dynamic LAIDS models. The forecasting experiment was designed to estimate and update static and dynamic models and subsequently generate one, two, three and four quarter-ahead forecasts for budget shares and consumption. The initial estimation period was from 1975(1) to 1999(4), and was updated one-observation-at-a-time up to 2002(3). Therefore, the forecast evaluation was conducted ex-post as was done in previous work. The ex-post forecast evaluation deviates, however, from most previous work in that tests of mean squared errors (MSE) are used to measure predictive ability of static and dynamic LAIDS models. The motivating reason for using MSE-tests is that a comparison of minimum MSE values many not reveal significant statistical differences between two competing models. In other words, a model with a smaller MSE may not provide more reliable data than its competitor when such differences are relatively small. The analysis found that the ECM-LAIDS model was superior in regards to forecasting performance when compared to the static model for beef at all four forecast horizons. In the case of pork, the static model performed better for one and two quarter-ahead forecasts when comparing minimum MSEs. Using MSE-tests, however, only the two-quarter-ahead forecasts were significant and in favor of the static model. The superior forecasting performance of the dynamic LAIDS models found in this study is consistent with findings in other applications (e.g., Chambers, 1997; Kastens & Brester, 1996; and Wang & Bessler, 2003).

The third case study in Chapter Five addressed the question of whether a more flexible modeling approach, a semiparametric model of LAIDS, would generate improved estimates of price and expenditure elasticities in cross-sectional data. The data were obtained from a 2003 Consumer Expenditure Survey (CES); this is a frequently used data source in cross-sectional estimations of demand systems. Although seemingly unrelated to the first two case studies, this case study analyzed another data property, particularly nonlinearity between budget shares and demographic variables, which are periodically discussed in the literature as a source of misspecification in demand systems. The chapter adopted Heckman's two step procedure to estimate the LAIDS model. The first set of estimates were parametric estimates from probit models of purchasing decisions and a seemingly unrelated regression (SUR) model that accounted for selectivity bias (also known as a parametric censored demand system). The second set of estimates came from a semiparametric specification of the censored demand systems, where age and household size entered both models, the probit equation and seemingly unrelated model, nonparametrically. Flexible seimparametric additive models were used to estimate the probit and SUR models. Initial exploratory analysis of the nonparametric relationship between budget shares and age and household size suggested a nonlinear relationship among them, consistent with nonlinearities periodically reported in previous work. In the final analysis, however, the semiparametric censored LAIDS model generated elasticity estimates that were qualitatively equivalent to those obtained from a parametric model. Unquestionably, the econometric literature on the semiparametric estimation of multivariate regression models is flourishing. As new developments are introduced into the literature, more generalized approaches to the semiparametric estimation of censored demand systems should be reassessed. Perhaps the merits of such promising modeling approaches can also be evaluated in other applications of demand systems.

REFERENCES

Anderson, F., and Blundell, R. “Testing restrictions in a flexible demand system: an application to consumer’s expenditure in Canada.” Review of Economic Studies 50(1983):397-410.

Ashley R., Granger C.W.J., and Schmalensee R. “Advertising and aggregate consumption: An Analysis of Causality.” Econometrica 48(1980):1149-1168.

Attfield, C.L.F. “Estimating a cointegrating demand system.” European Economic Review 41(1997):61-73.

Bachmeier, L. and Li, Q. “Is the term structure non-linear? A semiparametric investigation”. Applied Economics Letters 9(2002):151-151.

Balcombe, K.G., and Davis, J.R. “An application of cointegration theory in the estimation of the almost ideal demand system for food consumption in Bulgaria.” Agricultural Economics 15(1996):47-60.

Banks, J., Blundell, R., and Lewbel, A. “Quadratic Engel curves and consumer demand.” Review of Economics and Statistics 79(1997):527-539.

Banerjee, A., Dolado, J., and Smith, G. “Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence.” Oxford Bulletin of Economics and Statistics 48(1986):253-278.

Barten A.P. "Maximum likelihood estimation of a complete system of demand equations." European Economic Review 1(1969): 7-73.

Barten, A. P. "Consumer demand functions under conditions of almost additive preferences." Econometrica 32(1964): 1-38.

Bhargava, A. “Identification and panel data models with endogenous regressors.” Review of Economic Studies 58(1991):129-140.

Bierens, H., and Pott-Buter H. A. “Specification of household expenditure functions and equivalence scales by nonparametric regression.” Econometric Reviews 9(1990):123-210.

Blundell, R., Duncan, A., and Pendakur, K. “Semiparametric estimation of consumer demand.” Journal of Applied Econometrics 13(1998):435-461.

Byrne, P., Capps, O., and Saha, A. "Analysis of food-away from home expenditure patterns for U.S. households, 1982-89." American Journal of Agricultural Economics 78(1996):614-627.

Cleveland, W.S. "Robust locally weighted regression and smoothing scatterplots." Journal of the American Statistical Association, 74(1979): 829-836.

Cleveland, W.S. and Devlin, S.J. "Locally weighted regression: An approach to regression analysis by local fitting." Journal of the American Statistical Association, 83(1988): 596-610.

Chalfant, J.A. "A globally flexible, almost ideal demand system." Journal of Business and Economic Statistics 5(1987):233-242.

Chambers, M.J. "Forecasting with demand systems." Journal of Econometrics 4(1990):363-376.

Chambers, M.J., and K.B. Nowman "Forecasting with the almost ideal demand system: Evidence from some alternative dynamic specifications." Applied Economics 29(1997):935- 943.

Chavas, J.P. "Structural change in the demand for meat." American Journal of Agricultural Economics 65(1983): 148-154.

Christensen, L.R., Jorgenson, D.W., and Lawrence J.L. "Transcendental logarithmic utility functions." American Economic Review 65(1975):367-383.

Deaton, A. and Muellbauer, J. "Economics and Consumer Behavior". Cambridge, Cambridge University Press (1993)

Deaton, A., and Muellbauer, J. "An almost ideal demand system." The American Economic Review 70(1980):312-326.

Dhar, T., and Foltz, J.D. "Milk by any other name ... consumer benefits from labeled milk." American Journal of Agricultural Economics 87(2005):214-228.

Dickey, D.A., and Fuller, W.A., 1981. "Likelihood ratio statistics for autoregressive time series with a unit root." Econometrica 49(1981):1057–1072.

Diebold, F. X. (1998). Elements of Forecasting. Cincinnati, OH: South-Western College.

Diebold, F.X., and Mariano, R.S. "Comparing predictive accuracy" Journal of Business and Economic Statistics 13(1995):253-263.

Dong, D., Shonkwiler, J.S., and Capps, O. "Estimation of demand functions using cross-sectional household data: The problem revisited." American Journal of Agricultural Economics 80(1998):466–73.

Dong, D., Gould, B.W., and Kaiser, H.M. "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.

Duffy, M. "Advertising and food, drink and tobacco consumption in the United Kingdom: A dynamic demand system." Agricultural Economics 1637(2002):1-20.

Eales, J.S., and Unnevehr, L.J. "Demand for beef and chicken products: Separability and structural change." American Journal of Agricultural Economics 70(1988):521-532.

Eales, J.S., and Unnevehr, L.J. "Simultaneity and structural change in U.S. meat demand." American Journal of Agricultural Economics 75(1993):259-268.

Engle, R.F., and Granger, C.W.J. "Cointegration and error correction: representation, estimation and testing." Econometrica 55(1987):251-276.

Fraser, I., and Moosa, I.A. "Demand Estimation in the presence of stochastic trend and seasonality: The case of meat demand in the United Kingdom." American Journal of Agricultural Economics 84(2002):83-89.

Frisch, R. "Annual survey of general economic theory: the problem of index numbers." Econometrica 4(1936):1-39.

Ghysels, E., Lee, H.S., and Noh, J. "Testing for unit roots in time series: some theoretical extensions and a Monte Carlo investigation." Journal of Econometrics. 62(1994):415–442.

Gao, X. M. and Shonkwiler, J.S. "Characterizing taste change in a model of U.S. meat demand: Correcting for spurious regression and measurement errors." Review of Agricultural Economics 15(1993): 313-324.

Gao, X., and Spreen, T. "A microeconometric analysis of the U.S. meat demand." Canadian Journal of Agricultural Economics 42(1994):397-412.

Green, R., and Alston, J.M. "Elasticities in AIDS models." American Journal of Agricultural Economics 72(1990):442-445.

Haines, P. S., Guilkey, D.K., and Popkin, B.M. “Modeling food consumption decisions as a two-step process.” American Journal of Agricultural Economics 70(1988):543-52.

Heckman, J.J. “Sample selection bias as a specification error.” Econometrica 47 (1979):153-162.

Heien, D., and Wessells, C.R. “Demand systems estimation with microdata: A Censored Regression Approach,” Journal of Business & Economic Statistics 8 (1990):365-371.

Huang, K.S., and Haidacher, R. “Estimation of a composite food demand system for the United States.” Journal of Business and Economic Statistics 1(1983):285-291.

Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S. “Seasonal integration and cointegration.” Journal of Econometrics 44(1990):215–238.

Johansen, S. “A statistical analysis of cointegrating vectors.” Journal of Economic Dynamics and Control 12(1988):231-254.

Karagiannis, G., Katranidis, S., and Velentzas, K. “An error correction almost ideal demand system for meat in Greece.” Agricultural Economics 22(2000):29-35.

Karagiannis, G., and Mergos, G.J. “Estimating theoretically consistent demand systems using cointegration techniques with application to Greek food data.” Economics Letters 74(2002):37-43.

Kastens, T.L., and Brester, G.W. “Model selection and forecasting ability of theory-constrained food demand systems.” American Journal of Agricultural Economics 78(1996):301-313.

Kesavan, T., Hassan, Z.A., Hensen, H.H., and S. R. Johnson. “A dynamic and long-run Structure in U.S. meat demand.” Canadian Journal of Agricultural Economics 41(1993): 139-53.

Kinsey, J. “Diverse demographics drive the food industry.” Choices 5(1990):22.

Kremers, J., Ericsson, N., and Dolado, J. “The power of cointegration tests.” Oxford Bulletin of Economics and Statistics 54(1992):325-348.

Ljung, G. and Box, G. “On a measure of lack of fit in time series models.” Biometrika 67(1978): 297-303.

Moschini, G., and Daniele, M. “Maintaining and testing separability in demand systems.” American Journal of Agricultural Economics 76(1994):61-74.

Nayga, Jr., R.M. “Sample selectivity models for away from home expenditures on wine and beer.” Applied Economics 28(1996):1421-1425.

Nayga, R.M. “Microdata expenditure analysis of disaggregate meat products.” Review of Agricultural Economics 17(1995):275-285.

Ng, S. “Testing for homogeneity in demand systems when the regressors are nonstationary.” Journal of Applied Econometrics 10(1995):147-163.

Osborn, D.R., Heravi, S., & Birchenhall, C.R. Seasonality and the order of integration for consumption, Oxford Bulletin of Economics and Statistics, 50(1988):361-377.

Park, T. “Forecast evaluation for multivariate time-series models: The U.S. cattle market.” Western Journal of Agricultural Economics 15(1990):133-143.

Park, J.L., and Rodney, H.B. “A demand systems analysis of food commodities by U.S. households segmented by income.” American Journal of Agricultural Economics 78(1996):290-301.

Pendakur, K. “Estimates and tests of base-independent equivalence scales.” Journal of Econometrics 88(1999):1-40.

Perali, F.P., and Chavas, J.P. “Estimation of censored demand equations.” American Journal of Agricultural Economics 82(2000):1022-37.

Perron, P.P. “Trends and random walks in macroeconomic time series: Further evidence and implications.” Journal of Economic Dynamics Control 12(1988):279-332.

Phillips, P.P. “Time series regression with a unit root.” Econometrica 55(1987):277-301.

Piggott, N.E. “The nested piglog model: An application to US food demand.” American Journal of Agricultural Economics 85(2003):1-15.

Piggott, N.E., and Marsh, T.L. “Does food safety information impact US meat demand?” American Journal of Agricultural Economics 86(2004):154-174.

Pope, R., Green, R., and Eales, J. “Testing for homogeneity and habit formation in a flexible demand specification of U.S. Meat Consumption.” American Journal of Agricultural Economics 62(1980): 778-785.

Raper, K. C., M. N. Wanzala, and R. M. Nayga. "Food expenditures and household demographic composition in the US: A demand systems approach." Applied Economics 34(2002):981-992.

Robinson, P. "Root-N-consistent semiparametric regression." Econometrica 56(1988): 931-954.

Robledo, C.W., and H.O. Zapata, and M. McCracken "Measuring predictive accuracy in agribusiness forecasting." Journal of American Academy of Business. October 2002.

Roy, R. "De l'utilite, contribution a la theorie des choix." Paris: Hermann 1942.

Senauer, B., Asp, E. and Kinsey, J. Food Trends and the Changing Consumer. Eagan Press, St. Paul, Minnesota. Statistics 1991.

Shephard, R.W. Cost and Production Functions (Princeton: Princeton University Press) (1953)

Stone, J.R.N. "Linear expenditure systems and demand analysis: An application to the pattern of British demand." Economic Journal 64(1954):511-527.

Theil, H. "The information approach to demand analysis." Econometrica 33(1965):67-87.

Theil, H., and Clements, W.K. Applied demand analysis results from system-wide approaches. Cambridge, MA Ballinger, 1987.

Tobin, J. "Liquidity preference as behavior towards risk," The Review of Economic Studies, 25(1958):65–86.

United States Bureau of Labor Statistics. Consumer Expenditure Survey, Diary Survey 2003.

United States Department of Agriculture, Economic Research Service. Red Meats Yearbook. Statistical Bulletin No. 921, Washington DC, 1995.

United States Department of Agriculture, Economic Research Service. Poultry Yearbook. Statistical Bulletin No. 916, Washington DC, 1995.

United States Department of Labor. Detailed Monthly Consumer Price Indices for 2003. Bureau of Labor Statistics 2003.

Wales T.J. and Woodland A.D. "Estimation of consumer demand systems with binding non-negativity constraints." Journal of Econometrics 21(1983):263-85.

Wang, Z., and Bessler, D.A. “Forecast evaluations in meat demand analysis.” Agribusiness 19(2003):505-524.

West, K.D. “Asymptotic inference about predictive ability.” Econometrica 64(1996):1067-1084.

West, K.D., and McCracken, M. W. “Regression-based tests of predictive ability.” International Economic Review 39(1998):817-840.

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

Yen, S.T., Lin, B.H., and Smallwood, D.M. “Quasi and simulated-likelihood approaches to censored demand systems: Food consumption by food stamp recipients in the United States.” American Journal of Agricultural Economics 85(2003):458-91.

VITA

Anil Kumar Sulgham completed his secondary education at St. Mary's Centenary College 1994, Hyderabad. He enrolled in the College of Agriculture, ANGRAU, India, and received the degree of Agricultural 1998. He attended University of Georgia, US, and earned a Master of Science degree in Agricultural Economics in 2002. In the Spring of 2003, he enrolled in the Agricultural Economics doctoral program at Louisiana State University. He is currently a candidate for the degree of Doctor of Philosophy, which will be conferred in December of 2006. Upon completion of his doctoral studies, he will be employed as Statistical Analyst in private firm JP-Morgan Chase.