THEORETICALLY VALID AGGREGATES IN THE ABSENCE OF HOMOTHETIC PREFERENCES, SEPARABLE UTILITY, AND COMPLETE PRICE DATA

Edward Charles Van Eenoo, Jr.

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

MASTERS OF SCIENCE
in
Agricultural and Applied Economics

Dr. Everett Peterson, Chair*
Dr. Anya McGuirk
Dr. Jeffery Alwang

September 18, 1998
Blacksburg, Virginia

Keywords: Aggregation, Homotheticity, Separability, Price indices, Generalized Composite Commodity Theorem, Translog, Meat demand

* Dr. Paul J. Driscoll served as Chair until his death September 8, 1997. The ideas presented in this thesis are largely attributable to his intellect and inspiration.

THEORETICALLY VALID AGGREGATES IN THE ABSENCE OF HOMOTHETIC PREFERENCES, SEPARABLE UTILITY, AND COMPLETE PRICE DATA

Edward Charles Van Eenoo, Jr.

ABSTRACT

The improper aggregation of commodities can have important consequences when estimating a system of group demand equations. Generally, aggregates are created under the assumptions that intra-group preferences are homothetic and the consumer's utility function is weakly separable over some partition. These assumptions place severe restrictions on the model that can significantly impact parameter and elasticity estimates. An alternative to imposing weak separability is to employ the Generalized Composite Commodity Theorem, which requires the relative intra-group commodity prices to be independent of the group price index. This study compares the results of estimating a demand system for composite beef, pork, and poultry products under the assumptions of weak separability and the Generalized Composite Commodity Theorem. Another important issue related to aggregation is the specification of an appropriate group price index. Price indices consistent with linear homogeneous preferences (a subset of the homothetic class of preferences) and non-homothetic intra-group preferences are identified and it is shown that several of the commonly employed indices are biased in the absence of complete price data.

ACKNOWLEDGMENTS

This thesis would have had no beginning without the inspiration of Dr. Paul Driscoll and would have had no end without the dedication of Dr. Everett Peterson. I am deeply indebted to them both. In addition, Drs. Anya McGuirk and Jeffrey Alwang deserve substantial credit for their many helpful comments. Many thanks to you all.

In fond memory of:

Edward Charles Van Eenoo, Sr.

Ray M. Piddington

Paul J. Driscoll

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1: INTRODUCTION

Neoclassical demand theory hypothesizes that consumers maximize utility by making rational consumption choices. The range of potential consumption choices is vast. Everything from essential goods such as food, clothing, and shelter to luxury items such as perfume, automobiles, and pets are included in the consumer's utility function. Although this expansive array of consumables poses no theoretical difficulties, empirical demand models must be reduced in scale to make them tractable. Therefore, aggregation of seemingly related commodities into a smaller number of composite groups is common in empirical demand analysis.

The creation of *theoretically valid* aggregates poses several barriers to the empirical researcher. One barrier, of particular relevance to analysis involving cross sectional data, occurs when price data are incomplete (i.e. not all commodity prices are observed). When a consumer fails to purchase an individual commodity, it is not possible to calculate price in the usual manner of dividing expenditure by quantity. Under these circumstances many of the commonly employed price indices yield biased group prices (i.e. relative price differences between composite groups are not accurately captured by the selected index). Until now, the importance of using unbiased index numbers when creating aggregates in the absence of complete price data has been ignored in the demand analysis literature.

Another major difficulty in developing aggregate demand models is the need to make broad assumptions about the nature of consumer preferences and the structure of the consumer's utility function. It is common in demand research to assume that intra-group preferences are homothetic and that the consumer's utility function is weakly separable for some partition of the data. Under these conditions, aggregate demands can be determined as functions of *group* prices and income alone. However, separability severely restricts potential substitution effects between *inter*-group commodities and homotheticity severely restricts the allocation of group expenditures among *intra*-group commodities. These restrictions are often imposed on aggregate demand models without testing their validity due to the lack of any tractable alternative.

Some work has been done towards developing less restrictive models of aggregate demand. Segerson and Mount (1985) develop an aggregate demand model using an

Almost Ideal Demand System specification for intra-group share equations that allows group prices to be defined using *two* indices. This approach is useful because it allows for *group* demands to be estimated without imposing the restrictive assumption of homothetic intra-group preferences. On the downside, weak separability of the utility function is still required and defining group prices with two indices makes *conventional* elasticity estimation impossible.

Lewbel (1996) generalizes the Hicks – Leontief Composite Commodity Theorem to create aggregates without separable utility. This approach to aggregation places restrictions on the relative prices of intra-group commodities but does not require homothetic intra-group preferences in order to estimate group demand equations. The difficulty with this approach lies in finding aggregates that satisfy the necessary price restrictions. Aggregation research is important because theoretically valid aggregates are fundamental to a properly specified aggregate demand model. The inappropriate aggregation of commodities can bias parameter and elasticity estimates.

The purpose of this study is to estimate a theoretically consistent aggregate demand model that requires the fewest possible separability assumptions and that imposes no implicit or explicit restrictions on intra-group consumer preferences.[1] Chapter 2 presents a basic review of the neoclassical consumer demand model and gives necessary and sufficient conditions for the creation of theoretically valid aggregates using three alternative methods. Price indices consistent with both linear homogeneous preferences (a subset of the homothetic class of preferences) and non-homothetic preferences for intra-group commodities are identified in Chapter 3.[2] It is also shown that the commonly invoked Ideal, Laspeyres, and Törnqvist – Theil price indices are biased in the absence of complete price data. Using cross sectional data on meat consumption from the National Livestock and Meat Board, demand equations for beef, pork, and poultry are estimated in Chapter 4. Income and price elasticities are calculated for each group under both the

[1] The model estimated in Chapter 4 only considers consumer demand for beef, pork, and poultry. Reducing the model to this level does require assuming weak separability of the meat products comprising these groups from all other goods in the consumer's utility function. However, aggregation of the individual meat products into broad beef, pork, and poultry groups was accomplished without making any further separability assumptions. The model in Chapter 4 can be expanded to include other commodity groups, as data become available.

assumptions of linear homogeneous and non-homothetic preferences for intra-group commodities. Unfortunately, the models could only be estimated for a small partition of the data and the results were inconclusive about the effects that these assumptions have on the demand model.

[2] The stronger restriction of linear homogeneity was required because a price index that was consistent with homothetic intra-group preferences *and* unbiased in the absence of complete price data could not be identified.

CHAPTER 2: MODELS OF AGGREGATE DEMAND

A common area of research in applied microeconomics is the estimation of consumer demand models. This research is important because the resulting parameter estimates can be used to calculate expenditure and price elasticities, which provide objective information about the nature of commodity markets. The standard, neoclassical model of consumer demand, presented in Section 2.1, yields demand equations for *individual commodities* as functions of *commodity* prices and expenditure through the maximization of a *disaggregate* utility function. Under certain conditions, the standard model can be modified to yield *group* demand equations as functions of *group* prices and expenditure through the maximization of an *aggregate* utility function. For example, Figure 2.1 illustrates various levels of commodity aggregation. At the top level, commodities are partitioned into broad groups such as shelter, clothing, food, and medical expenses. These broad groups can then be partitioned into various sub-groups. For instance, the food group can be partitioned into cereals, dairy, meats, and produce. Commodities in the meat group can be more narrowly divided into beef, poultry, pork, and fish. This process continues until the lowest level is reached where the groups are defined as individual commodities (i.e. steak).[3] The demand equations for the individual commodities comprising the lowest level of the utility tree can be derived using relatively few and reasonable assumptions.[4] However, the demand equations for the aggregated commodities located at higher levels of the utility tree can only be estimated when much stricter restrictions are placed either on relative prices or the structure of the consumer's utility function *and* the nature of consumer preferences. Sections 2.2 and 2.3 explore the necessary and sufficient conditions for the existence of group demand equations and discuss the limitations imposed by these restrictions.

[3] Obviously, steak could be further divided into specific varieties such as T-bone, New York Strip, etc.
[4] These assumptions are discussed in more detail in Section 2.1

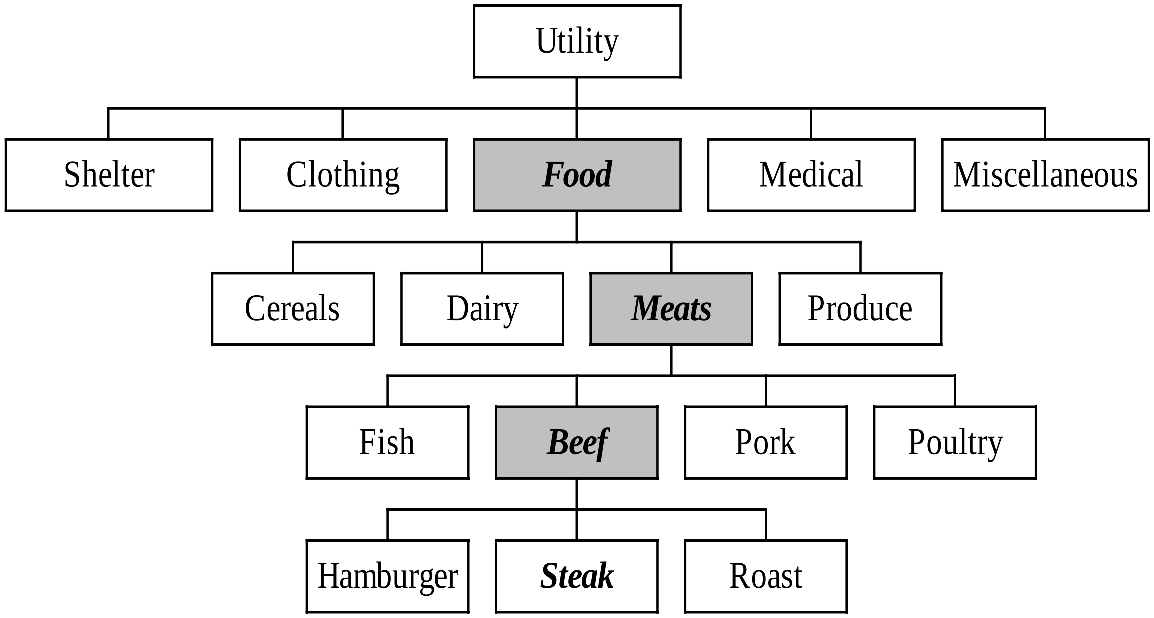

Figure 2.1: Various Levels of Commodity Aggregation

Note: This figure details various levels of commodity aggregation. The top level represents broad groups of commodities with common characteristics. These groups are more narrowly defined in each of the lower levels. The “groups” at the bottom level are defined as individual commodities.

2.1 Neoclassical Consumer Theory

In the standard neoclassical model of consumer demand, individuals are hypothesized to maximize utility by choosing from a finite set of consumption bundles subject to exogenously determined prices and a fixed level of expenditure. Under the assumptions that consumer preferences are reflexive, transitive, complete, and continuous a unique ordering of the consumption bundles can be represented by a continuous utility function . Formally, consumers

$$(2.1) \qquad \text{maximize}\ \{U(\mathbf{x}) \mid \mathbf{p} \cdot \mathbf{x} \leq y; x_i \geq 0, i = 1, \ldots, n\},$$

where **x** is an *n*-vector of commodities, **p** is a corresponding vector of commodity prices, and **p** · **x** = y is income.[5] With the additional assumption of local nonsatiation, the budget constraint must be met with equality and the consumer's problem can be restated as

$$(2.2) \qquad \text{maximize}\ \{U(\mathbf{x}) \mid \mathbf{p} \cdot \mathbf{x} = y; x_i \geq 0, i = 1, \ldots, n\}$$

The Lagrangian function associated with (2.2) is

$$(2.3) \qquad L = U(\mathbf{x}) + \lambda(y - \mathbf{p} \cdot \mathbf{x})$$

and the first-order necessary conditions for a maximum are

$$(2.4) \qquad \frac{\partial L}{\partial x_i} = \frac{\partial U(\mathbf{x})}{\partial x_i} - \lambda p_i = 0$$

$$(2.5) \qquad \frac{\partial L}{\partial \lambda} = y - \mathbf{p} \cdot \mathbf{x} = 0.$$

Solving the first order conditions for x_i yields ordinary demand equations of the form $x_i(\mathbf{p}, y)$. Consumer demands for individual commodities in the consumption bundle are functions of commodity prices and income.

The second-order sufficient condition for a maximum is negative definiteness of the bordered hessian

[5] The notation **(**·**)** represents the inner product of two vectors.

$$\mathbf{H} = \begin{pmatrix} 0 & g_1 & g_2 & \cdots & g_j \\ g_1 & L_{11} & L_{12} & \cdots & L_{1j} \\ g_2 & L_{21} & L_{22} & \cdots & L_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_i & L_{i1} & L_{i2} & \cdots & L_{ij} \end{pmatrix}$$

where g_i are the partial derivatives of the constraint equation and L_{ij} are the second order partial derivatives of (2.3):

$$g_i = \frac{\partial y}{\partial x_i}, \qquad L_{ij} = \frac{\partial^2 L}{\partial x_i \, \partial x_j} .$$

Dual to the direct utility function defined in (2.2) is the indirect utility function given by, Empirically, an indirect utility model is convenient because demand equations can be

(2.6) $$V(\mathbf{p}, y) = U(\mathbf{x}(\mathbf{p}, y)).$$

obtained directly through Roy's identity[6] and are generally of closed form. The indirect utility function defined by (2.6) can potentially be of large dimension causing degrees of freedom to be insufficient for estimation. For instance, assume **p** consists of *n* elements. Then (2.6) can be written as

(2.7) $$V(\mathbf{p}, y) = V(p_1, p_2, \ldots, p_n, y)$$

and the demand equations for individual commodities derived with Roy's identity are of the form,

(2.8) $$x_i(\mathbf{p}, y) = x_i(p_1, p_2, \ldots, p_n, y)$$

---

[6] Roy's identity is given by $x_i(\mathbf{p}, y) = -\dfrac{\dfrac{\partial V(\mathbf{p}, y)}{\partial p_i}}{\dfrac{\partial V(\mathbf{p}, y)}{\partial y}}$.

for $i = 1, 2, \ldots, n$. Demand equations for individual commodities are given as functions of the prices of *all* goods in the consumption vector and income. Since it is not generally possible to estimate the parameters of (2.8) when n is large, aggregation of commodities into a smaller number of seemingly related groups is necessary to reduce the dimensions of the model to a manageable level.

2.2 Aggregation with Separable Utility and Homothetic Preferences

The indirect utility function given in (2.7) implicitly assumes that individuals make consumption choices simultaneously, dependent upon the prices of *all* goods in the consumption bundle and income. This *one stage* budgeting procedure is theoretically desirable because it imposes the fewest possible restrictions on consumer preferences and captures all potential substitution effects between goods. However, when degrees of freedom are insufficient for estimation it is necessary to reduce the scale of the model. One means of accomplishing this is to consider the budgeting procedure as occurring in two (or more) stages.

In the first stage, of a *two stage* budgeting procedure, consumers allocate total expenditures among broad groups of commodities. In the second stage consumers allocate group expenditures among individual commodities. If two stage budgeting accurately represents a consumer's decision making process then the number of variables in the utility model can be drastically reduced without affecting the utility maximizing allocation of income. A necessary and sufficient condition for the *second stage* of a two stage budgeting procedure is weak separability of the utility function.

Consider partitioning the consumption vector **x** of (2.2) into m mutually exclusive subvectors $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}$ such that $\mathbf{x} = \mathbf{x_1} \cup \mathbf{x_2} \cup \ldots \cup \mathbf{x_m}$ and m < n. If weak separability holds under this partition then the utility function in (2.2) can be written as

$$(2.9) \qquad U(\mathbf{x}) = U[u_1(\mathbf{x_1}), u_2(\mathbf{x_2}), \ldots, u_m(\mathbf{x_m})]$$

where each u_I is a strictly monotonic, quasi-concave subutility (or aggregator) function for $I = 1, 2, \ldots, m$. Now, the utility maximizing level of each x_i can be determined by maximizing the relevant subutility function subject to expenditures on that group.

This specification is useful because it allows the demand for commodities in one group to be analyzed independently of the quantity of commodities consumed in other groups. For example, let $u_I(\mathbf{x_I})$ denote the level of utility derived from consumption of commodities in group *I* and let $m_I = \mathbf{p_I} \cdot \mathbf{x_I}$ denote total expenditures on group *I*, where the vector $\mathbf{x_I}$ is defined as

$$\mathbf{x_I} = \left[(x_i) \middle| \, i \in I \right]$$

and $\mathbf{p_I}$ is a corresponding vector of prices. If weak separability holds then individual demand equations for the commodities comprising group *I* will be of the form

$$x_{i \in I} = x_i \left(\mathbf{p_I}, m_I \right) \tag{2.10}$$

for $i = 1, 2, \ldots, n$.[7] Compare these results to the disaggregate model defined by (2.7) where demand equations are functions of the prices of *all (n)* goods and total expenditure (see (2.8)). Given weak separability of the utility function in (2.2), under the partition defined by $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}$, individual demands for the various commodities are now given solely as functions of intra-group commodity prices and group expenditures.

Weak separability is *not* a sufficient condition for the *first stage* of a two stage budgeting procedure. To estimate the demand for each composite group $I = 1, \ldots, m$ as functions of group prices and income only, additional restrictions must be placed upon the model. In particular, each aggregator function, u_I (see equation (2.9)) must be homothetic [Gorman (1959)].

When the aggregator functions are restricted to homothetic forms, a unique solution to the consumers problem

$$\text{maximize} \left\{ U\left(u_1(\mathbf{x_1}), u_2(\mathbf{x_2}), \ldots, u_m(\mathbf{x_m})\right) \middle| \sum_{I=1}^{m} \mathbf{p_I} \cdot \mathbf{x_I} = y; x_i \geq 0 \; \forall \, i \right\} \tag{2.11}$$

can be found as a function of group prices and income. Mathematically, group demands arising from the solution to (2.11) are of the form

$$X_I = X_I \left(P_1, P_2, \ldots, P_m, y \right) \tag{2.12}$$

where X_I are group demands and P_I are group prices determined by some price index appropriate to the model and data.

Assumptions of weakly separable utility functions and homothetic aggregator functions are commonly invoked in empirical research to reduce the number of model variables and make parameter estimation possible. However, these assumptions place severe restrictions on the model. Homotheticity requires intra-group budget shares to be independent of total group expenditure and separability severely limits the potential substitution effects between inter-group commodities. These restrictions may be unrealistic and have not been supported by past empirical research [see, for example, Diewert and Wales (1995)].

2.3 Aggregation when Utility is Non-Separable and Preferences are Non-Homothetic

Until recently, the only alternative to separability for reducing the dimensions of empirical demand models was the Hicks – Leontief Composite Commodity Theorem. Instead of restricting intra-group preferences and inter-group substitution effects, theoretically valid aggregates can be created (and group demands estimated) by restricting intra-group relative price movements.

Once again, consider dividing the consumption vector **x** from (2.2) into m mutually exclusive subvectors $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}$ such that $\mathbf{x} = \mathbf{x_1} \cup \mathbf{x_2} \cup \ldots \cup \mathbf{x_m}$ and m < n. Let $\mathbf{p_1}, \mathbf{p_2}, \ldots, \mathbf{p_m}$ be the corresponding price subvectors. Using a price index appropriate to the data and model assumptions, define P_I as the group price for each group $I = 1, 2, \ldots, m$. Define ρ_i by

$$(2.13) \qquad \rho_{i \in I} = \log\left(\frac{p_i}{P_I}\right)$$

for $I = 1, 2, \ldots, m$. The Hicks – Leontief Theorem states that group demand equations can be determined as functions of group prices and expenditure if

$$\rho_i^r = \rho_i^s \text{ for } i = 1,2,\ldots,n \text{ and } r \neq s$$

[7] Deaton and Muellbauer (1980, p.124) give a simple proof of this result.

where the superscripts r and s refer to individual consumers and/or time periods. This is equivalent to saying that relative commodity prices, ρ_i, are constant across consumers. Although it seems likely for relative commodity prices to be strongly correlated between consumers, the Hicks – Leontief Composite Commodity Theorem unrealistically requires the correlation to equal one.

The Generalized Composite Commodity Theorem (GCC) [Lewbel (1996)] relaxes the restrictive assumption of constant relative prices imposed under Hicks – Leontief aggregation by allowing relative prices to vary across consumers. Following Lewbel, define $r_i = ln(p_i)$, $R_I = ln(P_I)$ and define ρ_i as in (2.13). Let $\mathbf{r_i}$, $\mathbf{R_I}$, and $\boldsymbol{\rho_i}$ represent the vectors of elements r_i, R_I and ρ_i. The GCC states that group demand equations, X_I, can be determined as a function of group prices and expenditure alone given the following conditions:

1. The individual commodity demand equations, x_i, are rational (i.e. consistent with the maximization of some utility function), and

2. The vectors of relative commodity prices, $\boldsymbol{\rho_i}$, are distributed independently of $\mathbf{R_I}$ and the logarithm of total expenditure for all i.

Notice that when relative prices, ρ_i, are constant across consumers, condition two must hold, implying that the Hicks – Leontief Composite Commodity Theorem is a special case of the GCC.

Although less restrictive than Hicks – Leontief or separability, the second condition of the GCC can be difficult to satisfy [Davis (1998)]. Statistical techniques such as cluster analysis [Nicol (1991)] can be helpful in identifying commodity groupings consistent with the GCC but the resulting groups may have no economic significance or relation to the problem being studied.[8] For instance, assume the problem is to estimate aggregate food demand but the group comprised of food products fails to satisfy the conditions of the GCC. Cluster analysis could be used to determine an appropriate grouping of commodities. However, this new grouping of commodities will be of little value in estimating food demand if the individual food commodities are spread out

[8] Davis (1998) uses cluster analysis to establish appropriate groupings of meat data beyond the standard beef, poultry, and pork categories. As expected, the groups identified included meats of different varieties and had little economic significance. The surprising conclusion of this study was that even the groups identified through cluster analysis did not satisfy the conditions of the GCC.

among various food and non-food groups. The same is true if non-food products are included in the group containing individual food commodities.

Given the need for economically meaningful groups, the choice between creating aggregates via separability versus the GCC may seem inconsequential. However, creating aggregates consistent with the GCC has several advantages. First, it is easy to determine when groups are inconsistent with the conditions of the GCC. It is well known that the independence of two random variables implies a correlation coefficient of zero. The contrapositive of this result gives a means for testing when condition 2 of the GCC is *not* satisfied. If the correlation coefficient between ρ_i and $\boldsymbol{R_I}$ is significantly different from zero then the two variables are dependent and the aggregation being conducted is not consistent with the GCC.

In contrast, separability is difficult to test for and the power of such tests is questionable [Alston and Chalfant (1991)]. Lewbel (1996) suggests that the inability to conclusively prove or disprove separability assumptions may be the reason for their ubiquitous use in the demand analysis literature and may have potentially led to the improper specification of many aggregate demand models. The consequences of inappropriately assuming separable utility are potentially serious and can lead to biased parameter and elasticity estimates.

A second advantage to creating aggregates consistent with GCC is the ability to use non-homothetic aggregator functions. When the aggregator functions are allowed to be non-homothetic, intra-group budget shares are free to vary with group expenditure. By plotting the Engel curve for two intra-group commodities, x_i and x_j, it is easy to demonstrate how non-homothetic aggregator functions can lead to a more general model of aggregate demand (see Figure 2.2). When the aggregator functions are homothetic, the Engel curve is restricted to a linear relationship between x_i and x_j. This means the allocation of group expenditure to x_i and x_j is not affected by changes in the level of group expenditure. This unrealistic restriction is avoided when non-homothetic aggregator functions are allowed, implying that budget shares can vary with group expenditure.

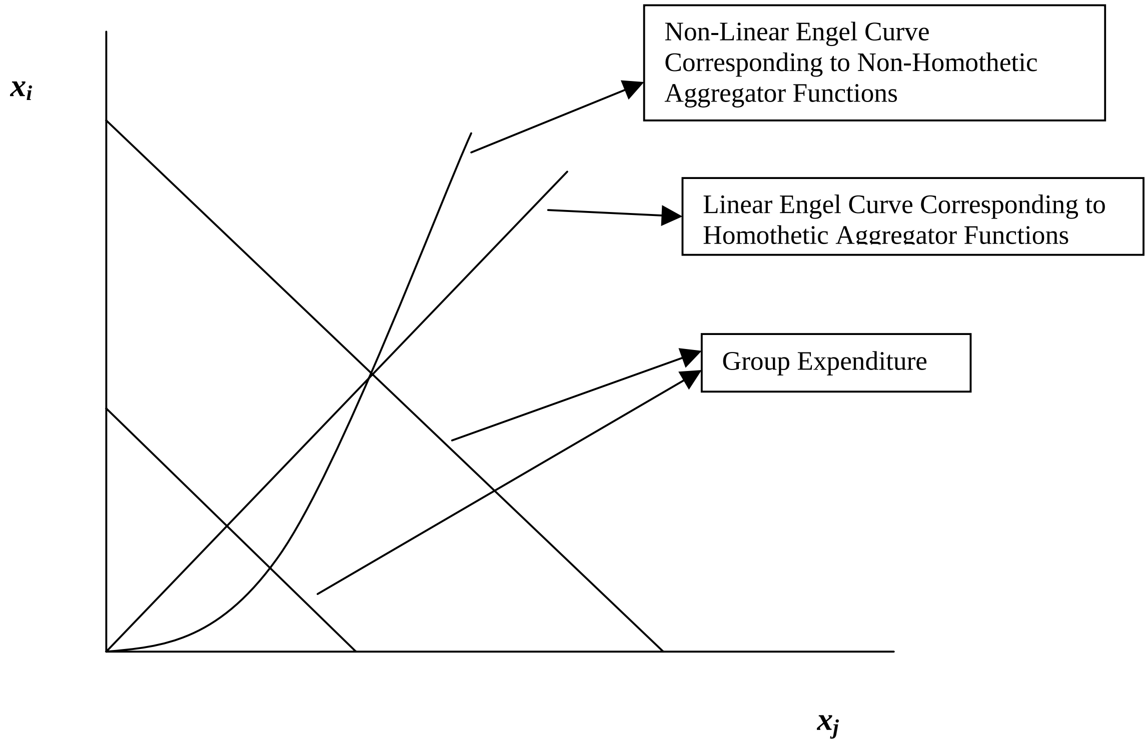

Figure 2.2: Engel Curves Corresponding to Homothetic and Non-Homothetic Aggregator Functions

Note: Homothetic preferences result in a Linear Engel curve, implying that the relative budget shares allocated to the consumption of x_i and x_j must remain constant regardless of expenditure. This restriction is avoided by allowing for non-homothetic preferences.

CHAPTER 3: THEORETICALLY VALID INDEX NUMBERS

The previous chapter demonstrated why the aggregation of commodities into a smaller number of seemingly related groups is often necessary in empirical demand analysis and presented three methods for creating theoretically valid aggregates. The purpose of this chapter is to establish appropriate price and quantity indices for these aggregate groups. Recall the partitioning of the consumption vector **x** defined by the aggregate utility model of (2.9): $\mathbf{x} = \mathbf{x_1} \cup \mathbf{x_2} \cup \ldots \cup \mathbf{x_m}$ where **x** is an n-vector of commodities, $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}$ are mutually exclusive, and m < n. Define P_I as the price index for group I and X_I as the corresponding quantity index for group I. Various methods exist for establishing P_I and X_I but the bilateral axiomatic approach is the most widely used. It is referred to as bilateral because P_I and X_I are defined as a ratio between some individual, or "observed" consumer[9] (time period), and some base consumer (time period). A useful result of the bilateral approach is Fischer's weak factor reversal condition, which requires P_I and X_I satisfy

$$(3.1) \qquad P_I \; X_I = \frac{\sum_{i \in I} p_i^1 x_i^1}{\sum_{i \in I} p_i^0 x_i^0}.$$

where p_i^1 refers to the price of commodity x_i^1 for the observed consumer and p_i^0 refers to the price of commodity x_i^0 for some base consumer. Given some arbitrary quantity index, X_I, weak factor reversal defines the corresponding price index, P_I, as

$$(3.2) \qquad P_I = \frac{\sum_{i \in I} p_i^1 x_i^1}{X_I \sum_{i \in I} p_i^0 x_i^0}.$$

Before giving examples of functional forms for P_I and X_I, it is necessary to further define the base, or "average," consumer. An arbitrary but reasonable approach for

[9] The consumer is referred to as "observed" because his consumption and expenditures have actually been observed instead of being imputed as in the case of the base consumer.

determining the consumption vector, $\mathbf{x^0}$, and a corresponding price vector, $\mathbf{p^0}$, for the base consumer is to define

$$\mathbf{x^0} \equiv \left(\frac{\sum_{s=1}^{S} x_1^s}{S}, \frac{\sum_{s=1}^{S} x_2^s}{S}, \ldots, \frac{\sum_{s=1}^{S} x_n^s}{S} \right), \text{ and}$$

$$\mathbf{p^0} \equiv \left(\frac{\sum_{s=1}^{S} m_1^s}{\sum_{s=1}^{S} x_1^s}, \frac{\sum_{s=1}^{S} m_2^s}{\sum_{s=1}^{S} x_1^s}, \ldots, \frac{\sum_{s=1}^{S} m_n^s}{\sum_{s=1}^{S} x_1^s} \right)$$

where s represents an individual consumer, S represents the total number of consumers, and m_i^s represents expenditures on commodity i by consumer s.[10] In words, $\mathbf{x^0}$ represents the average quantity consumed of each individual commodity by all consumers and $\mathbf{p^0}$ is a corresponding vector of average prices paid.

Define $\mathbf{x^1}$ and $\mathbf{p^1}$ as consumption and price vectors for the observed consumer. Examples of functional forms for P_I commonly found in the literature are:

$$(3.3) \qquad P_I^{id} \equiv \sqrt{\frac{(\mathbf{p_I^1} \cdot \mathbf{x_I^0})(\mathbf{p_I^1} \cdot \mathbf{x_I^1})}{(\mathbf{p_I^0} \cdot \mathbf{x_I^0})(\mathbf{p_I^0} \cdot \mathbf{x_I^1})}} \qquad [\text{Ideal Price Index}],$$

$$(3.4) \qquad P_I^{pa} \equiv \frac{(\mathbf{p_I^1} \cdot \mathbf{x_I^1})}{(\mathbf{p_I^0} \cdot \mathbf{x_I^1})} \qquad [\text{Paasche Price Index}],$$

$$(3.5) \qquad P_I^{la} \equiv \frac{(\mathbf{p_I^1} \cdot \mathbf{x_I^0})}{(\mathbf{p_I^0} \cdot \mathbf{x_I^0})} \qquad [\text{Laspeyres Price Index}], \text{ and}$$

$$(3.6) \qquad P_I^{tt} \equiv \prod_{i \in I} \left(\frac{p_i^1}{p_i^0} \right)^{.5\left(\frac{p_i^1 x_i^1}{\mathbf{p_I^1} \cdot \mathbf{x_I^1}} + \frac{p_i^0 x_i^0}{\mathbf{p_I^0} \cdot \mathbf{x_I^0}} \right)} \qquad [\text{Tornqvist - Theil Price Index}],$$

where the subscript I is understood to refer to a specific group and the underlying commodities that comprise that group for I = 1, 2, …, m. Functional forms for the group

[10] Implicit in this definition of $\mathbf{x^0}$ and $\mathbf{p^0}$ is the assumption that the commodities being purchased by the various consumers are homogenous. Where substantial brand differences exist (e.g. salad dressings), this definition may not be appropriate.

quantity analogues to (3.3) – (3.6) are found by exchanging quantities and prices in the four equations.

The variety of potential index numbers leads to ambiguity in the choice of P_I and X_I and is the main criticism of the axiomatic approach to index number theory [Frisch (1936)]. To reduce the ambiguity, the axiomatic approach relies upon a set of seemingly reasonable tests (or axioms) that are sufficient for determining a unique functional form for P_I and X_I. An exhaustive set of mutually consistent axioms has yet to be established that can yield a unique form for P_I and X_I. However, the existing axioms clearly indicate that some index numbers are better than others.

Section 3.1 lists nine axioms commonly found in the literature. The Ideal, Paasche, Laspeyres, and Törnqvist – Theil (TT) indices are tested against these axioms to determine which is the "best" index. The presence of unobservable commodity prices adds additional difficulties to the choice of an appropriate index number. These difficulties and potential solutions are discussed in Section 3.2. Finally, Section 3.3 proves that the Paasche index is consistent with the assumption of linear homogeneous aggregator functions and suggests the TT index as an appropriate choice when non-homothetic aggregator functions are assumed.

3.1 Axiomatic Approach to Index Number Theory

Diewert (1987) identifies nine axioms that are commonly used in the literature as benchmarks for measuring the quality of index numbers.[11] Diewert confines his analysis to time series data but the axioms are identical for cross sectional data. This section presents these nine axioms and provides explanations relevant to cross sectional data. In the following axioms, the subscript I is dropped to simplify notation and P refers to some arbitrary price index.

<u>*Axiom 1 (Identity):*</u> Assume $\alpha > 0$, $\beta > 0$, $\mathbf{p^0} = \mathbf{p^1}$ and $\mathbf{x^0} = \mathbf{x^1}$. Then

$$P\left(\mathbf{p^0}, \mathbf{p^1}, \alpha\mathbf{x^0}, \beta\mathbf{x^1}\right) = 1.$$

[11] The notation used in this section closely follows that of Diewert.

This simply states that if the base and observed consumer face the same prices and purchase each commodity in the same proportion, then the group price for the observed consumer should equal one.

Axiom 2 (Proportionality): Assume α > 0. Then

$$P(\mathbf{p^0}, \alpha\mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}) = \alpha P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

In other words, if each commodity price faced by the observed consumer increases α percent then the group price P should also increase α percent.

Axiom 3 (Invariance to Changes in Scale): Assume $\alpha > 0$, $\beta > 0$, and $\gamma > 0$. Then

$$P(\alpha\mathbf{p^0}, \alpha\mathbf{p^1}, \beta\mathbf{x^0}, \gamma\mathbf{x^1}) = P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

This axiom captures many properties desirable in an index number. First, if commodity prices faced by the base consumer *and* observed consumer change by the same percentage (α), then P should remain constant. This allows the scale of price measurement to change (i.e. from American Dollars to British Pounds) without altering group prices. Second, if the consumption of all commodities by the base or observed consumer increases by the same percentage, then group prices should remain constant.

Axiom 4 (Invariance to Changes in Units): Assume $\mathbf{p^i}$ and $\mathbf{x^i}$ are n-vectors for i = 0, 1. Define

$$\tilde{\alpha} \equiv (\alpha_1, \alpha_2, \ldots, \alpha_n), \text{ and}$$
$$\tilde{\alpha}^{-1} \equiv (\alpha_1^{-1}, \alpha_2^{-1}, \ldots, \alpha_n^{-1}).$$

If $\alpha_i > 0$ for $i = 1, 2, \ldots, n$ then

$$P(\tilde{\alpha} \cdot \mathbf{p^0}, \tilde{\alpha} \cdot \mathbf{p^1}, \tilde{\alpha}^{-1} \cdot \mathbf{x^0}, \tilde{\alpha}^{-1} \cdot \mathbf{x^1}) = P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

In words, if a commodity price increase (decrease) is offset by proportional decrease (increase) in the consumption of that commodity (i.e. consumption decreases enough to keep group expenditures constant), then group prices should remain constant. This allows for the scale of quantity measurement to change (i.e. from pounds to kilograms) without altering group prices.

<u>*Axiom 5 (Consumer Reversal):*</u> Assume that a group price, P, is established using some arbitrary price index. Reversing the price and consumption vectors for the base and observed consumer should result in the inverse of P. Mathematically,

$$P\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right) = \frac{1}{P\left(\mathbf{p^1}, \mathbf{p^0}, \mathbf{x^1}, \mathbf{x^0}\right)}.$$

A simple example can clarify this. Assume P = ½. This implies that the group price for the observed consumer is one-half that of the base consumer. It seems logical therefore for group price to equal 2 when the role of the base consumer and observed consumer are switched. That is P($\mathbf{p^0}$, $\mathbf{p^1}$, $\mathbf{x^0}$, $\mathbf{x^1}$) = 1/ P($\mathbf{p^1}$, $\mathbf{p^0}$, $\mathbf{x^1}$, $\mathbf{x^0}$)= 2.

<u>*Axiom 6 (Commodity Reversal):*</u> Define $\mathbf{p}^{\hat{\mathbf{i}}}$ as a permutation of $\mathbf{p^i}$ and $\mathbf{x}^{\hat{\mathbf{i}}}$ as the same permutation of $\mathbf{x^i}$ for i = 0, 1. Then

$$P\left(\mathbf{p}^{\hat{0}}, \mathbf{p}^{\hat{1}}, \mathbf{x}^{\hat{0}}, \mathbf{x}^{\hat{1}}\right) = P\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right).$$

For example, assume group I is comprised of only two commodities and that P is defined by the Paasche index (3.4). Then, if commodity reversal holds,

$$\text{(3.7)} \qquad \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^1}\right)} = \frac{\left(\mathbf{p_I^{\hat{1}}} \cdot \mathbf{x_I^{\hat{1}}}\right)}{\left(\mathbf{p_I^{\hat{0}}} \cdot \mathbf{x_I^{\hat{1}}}\right)}.$$

where

$$\mathbf{p_I^i} = (p_1^i, p_2^i),$$
$$\mathbf{p_I^{\hat{i}}} = (p_2^i, p_1^i),$$
$$\mathbf{x_I^i} = (x_1^i, x_2^i), \text{ and}$$
$$\mathbf{x_I^{\hat{i}}} = (x_2^i, x_1^i)$$

for $i = 0, 1$. Substituting these values into (3.7) yields

$$\frac{(p_1^1 x_1^1) + (p_2^1 x_2^1)}{(p_1^0 x_1^1) + (p_2^0 x_2^1)} = \frac{(p_2^1 x_2^1) + (p_1^1 x_1^1)}{(p_2^0 x_2^1) + (p_1^0 x_1^1)}$$

which is clearly true and commodity reversal is satisfied for this example.

Axiom 7 (Monotonicity): If at least one of the commodity prices faced by the observed consumer decreases (increases) then the corresponding group price must also decrease (increase), ceteris paribus. Mathematically,

$$P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}) \geq P(\mathbf{p^0}, \mathbf{p^{\hat{1}}}, \mathbf{x^0}, \mathbf{x^1}) \text{ where } p_i^{\hat{1}} \leq p_i^1 \ \forall\, i \in I.$$

Axiom 8 (Mean Value Test): Group prices must lie between the smallest and largest ratios of commodity prices between the observed and base consumers. Mathematically,

$$\min_i \left(\frac{p_i^1}{p_i^0} \right) \leq P \leq \max_i \left(\frac{p_i^1}{p_i^0} \right).$$

For example, assume $\mathbf{p^1} = (p_1^1, p_2^1, p_3^1) = (\$2.50, \$3.00, \$2.25)$ and $\mathbf{p^0} = (p_1^0, p_2^0, p_3^0) = (\$1.50, \$3.00, \$1.75)$. The Mean Value Test requires that P fall in the range

$$\min_i \left(\frac{p_i^1}{p_i^0} \right) = \left(\frac{p_2^1}{p_2^0} \right) = \left(\frac{\$3.00}{\$3.00} \right) = 1 \leq P \leq 1.67 = \left(\frac{\$2.50}{\$1.50} \right) = \left(\frac{p_1^1}{p_1^0} \right) = \max_i \left(\frac{p_i^1}{p_i^0} \right).$$

Axiom 9 (Circularity): Circularity is a transitivity condition that is more relevant to time series data. Define $\mathbf{p^2}$ as the price vector observed in period 2. Define $\mathbf{p^1}$ as the price vector observed in period 1. Define $\mathbf{p^0}$ as the price vector observed in the base period. Define $\mathbf{x^2}$, $\mathbf{x^1}$ and $\mathbf{x^0}$ as the consumption vectors corresponding to $\mathbf{p^2}$, $\mathbf{p^1}$ and $\mathbf{p^0}$ respectively. Circularity requires

$$P(\mathbf{p}^0, \mathbf{p}^1, \mathbf{x}^0, \mathbf{x}^1)P(\mathbf{p}^1, \mathbf{p}^2, \mathbf{x}^1, \mathbf{x}^2) = P(\mathbf{p}^0, \mathbf{p}^2, \mathbf{x}^0, \mathbf{x}^2).$$

In words, a price index derived between period 1 and the base period multiplied by a price index derived between period 2 and period 1 should equal a price index derived between period 2 and the base period directly.

The four indices defined in (3.3) – (3.6) are analyzed against these nine axioms in Appendix A. The Ideal index fails only Axiom 9, the Laspeyres and Paasche indices fail Axioms 5 and 9, and the TT index fails Axioms 7 and 9. Thus, based upon these nine axioms, the Ideal index would appear to be the best of the four. However, as the next section shows, the Ideal index yields biased aggregate prices when some of the commodity prices are unobserved.

3.2 Price Indices Appropriate in the Absence of Complete Price Data

A common problem in establishing group price P_I occurs when a consumer purchases zero quantity of one or more of the commodities belonging to group I. A similar problem arises in time series data when new products are introduced to a group. To understand the problem caused by unobserved prices, recall the price indices defined by (3.3) – (3.6). Each requires price data for all commodities for both the base consumer and observed consumer. However, in cross-sectional data, it is common not to have price information included for non-consumed commodities. In addition, when a commodity is not purchased, it is not possible to calculate the unobserved price in the usual manner of dividing expenditure by quantity. The remainder of this section discusses three potential solutions to the problem of unobserved prices.

The simplest solution is to ignore any commodities that were not purchased by the observed consumer. This approach was originally advocated by Keynes (1930) as a solution to the new good problem associated with time series data. However, ignoring commodities when price is unknown can bias group prices and the bias is likely to increase with the number of commodities ignored. For example, in time series data, there are generally few new commodities introduced during any period relative to the total number of commodities in that group. Thus, the bias introduced from ignoring these commodities in determining group price is likely to be small. On the other hand, when data are cross-sectional, it is quite common for a consumer to not purchase the majority of goods in a group. In this situation, the bias introduced from ignoring the commodities not purchased is likely to be large. Therefore, when data is cross-sectional, this approach is not recommended.

Hicks (1940) suggested using shadow (or reservation) prices in place of the unknown prices. In other words, the unknown prices are set equal to the exact price at which the consumer would have chosen to consume zero of the commodity. However, this approach is of little empirical value since the shadow prices themselves are not observable and the econometric techniques for estimating them are cumbersome [Diewert (1987)].

A third approach, previously not identified in the literature, is to use an index number that is not biased by incomplete price data. Recall the Paasche index given in (3.4). Assume group I consists of n-elements. Thus, the consumption vector for the observed consumer is given by $\mathbf{x_I^1} = (x_1^1, x_2^1, \ldots, x_n^1)$. The elements of the price vector $\mathbf{p_I^1} = (p_1^1, p_2^1, \ldots, p_n^1)$, corresponding to the consumption vector, will be indeterminate only when $x_i^1 = 0$. However, when $x_i^1 = 0$, the price corresponding to commodity i, p_i^1, can be set equal to any value without affecting the value of P_I^{pa} [see (3.4)]. Thus, the Paasche index is not biased in the absence of complete price data.

The Ideal and Laspeyres price indices are not appropriate when price data are incomplete because of the term $(\mathbf{p_I^1 . x_I^0})$. Since $\mathbf{x_I^0} = (x_1^0, x_2^0, \ldots, x_n^0)$ is strictly positive for all i, it is necessary to have accurate commodity prices for all p_i^1. Setting indeterminate elements of p_i^1 equal to some arbitrarily value can bias group prices defined by P_I^{id} and P_I^{la}. For example, assume group I is comprised of three commodities and that consumer 1 purchases the three commodities as defined by $\mathbf{x_I}^{\mathbf{1}} = (x_1^1, x_2^1, x_3^1) = (0, 1, 0)$. Let the actual vector of prices faced by consumer 1 be defined by $\mathbf{p_I}^{\mathbf{1}} = (p_1^1, p_2^1, p_3^1) = (\$1.00, \$1.50, \$2.00)$. For computational convenience, let $\mathbf{p_I}^{\mathbf{0}} = (p_1^0, p_2^0, p_3^0)$

= *($1.00, $1.00, $1.00)* and $\mathbf{x_I}^{\mathbf{0}}$ = $(x_1^0, x_2^0, x_3^0) = (1, 1, 1)$.[12] Using the Laspeyres price index defined in (3.5) yields

$$P_I^{la} \equiv \frac{(\mathbf{p_I^1} \cdot \mathbf{x_I^0})}{(\mathbf{p_I^0} \cdot \mathbf{x_I^0})} = \frac{\$4.50}{\$3.00} = 1.50$$

as the true group *I* price index for consumer 1. However, since x_1 and x_3 are not purchased, it is not possible to impute p_1^1 and p_3^1 and $\mathbf{p_I}^{\mathbf{1}}$ = *(unknown, $1.50, unknown)*. Choosing inappropriate values for p_1^1 and p_3^1 will bias P_I^{la}. For example, it may seem intuitive to set unobserved prices equal to zero yielding $\mathbf{p_I}^{\mathbf{1}}$ = *(0, $1.50, 0)*. This results in

$$P_I^{la} = \frac{\$1.50}{\$3.00} = 0.50$$

which is approximately 66% lower than the true price index for group *I*.

The TT index also yields biased prices when price data are incomplete. Using the above values for $\mathbf{p_I}^{\mathbf{0}}$, $\mathbf{p_I}^{\mathbf{1}}$, $\mathbf{x_I}^{\mathbf{0}}$, and $\mathbf{x_I}^{\mathbf{1}}$ results in a true index value of

$$P_I^{tt} \equiv \prod_{i \in I} \left(\frac{p_i^1}{p_i^0} \right)^{.5\left(\frac{p_i^1 x_i^1}{\mathbf{p^1} \cdot \mathbf{x^1}} + \frac{p_i^0 x_i^0}{\mathbf{p^0} \cdot \mathbf{x^0}} \right)} = 1.47 \, .$$

When zeros are inserted for the unknown price information, the TT index yields

$$P_I^{tt} \equiv \prod_{i \in I} \left(\frac{p_i^1}{p_i^0} \right)^{.5\left(\frac{p_i^1 x_i^1}{\mathbf{p^1} \cdot \mathbf{x^1}} + \frac{p_i^0 x_i^0}{\mathbf{p^0} \cdot \mathbf{x^0}} \right)} = 1.31$$

which is approximately 11% lower than the true index value.

Next, consider the same example using the Paasche price index defined in (3.4) for the price of group *I*. When the actual commodity prices are used, the Paasche index yields

[12] Note that this definition of $\mathbf{x_I^0}$ comes at no loss of generality. Any positive, non-zero values of x_1^0, x_2^0,

$$P_I^{pa} \equiv \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^1}\right)} = \frac{\$1.50}{\$1.00} = 1.50$$

When zeros are inserted for the missing price information, the Paasche price index yields

$$P_I^{pa} \equiv \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^1}\right)} = \frac{\$1.50}{\$1.00} = 1.50\,,$$

illustrating that the Paasche price index is unbiased by incomplete price data. Therefore, when price data for the observed consumer is incomplete, it is necessary to consider factors beyond the nine axioms given in Section 3.1 when choosing an appropriate price index to insure that the chosen index is not biased by unobserved prices. One final criterion that must be considered in choosing an index number is the assumed structure of the aggregator functions defined in (2.9).

3.3 Price Indices Consistent with Homothetic and Non-Homothetic Aggregator Functions

A generalized, weakly separable utility function can be defined by

$$(3.8) \qquad U(\mathbf{x}) = U\left[u_1(\mathbf{x_1}), u_2(\mathbf{x_2}), \ldots, u_m(\mathbf{x_m})\right]$$

where the consumption vector x = $x_1 \cup x_2 \cup \ldots \cup x_m$ and the subvectors $x_1, x_2, \ldots,$ x_m are mutually exclusive. To insure rationality of group demand functions it is necessary to assume homotheticity of u_I for $I = 1, 2, \ldots, m$ [see Section(2.2)]. A first order Taylor series expansion of u_I is

$$(3.9) \qquad u_I^1 = u_I^0(\mathbf{x_I^0}) + \sum_{i \in I} \frac{\partial u_I}{\partial x_i^0}\left(x_i^1 - x_i^0\right)$$

which implies

---

x_3^0 will produce similar results.

$$(3.10) \qquad u_I^1 = u_I^0\left(\mathbf{x_I^0}\right) + \sum_{i \in I} \frac{\partial u_I}{\partial x_i^0} x_i^1 - \sum_{i \in I} \frac{\partial u_I}{\partial x_i^0} x_i^0 \,.$$

Simplification of (3.10) requires Euler's theorem. Assume $f = f(x_1, x_2, \ldots, x_n)$ is a linear homogenous function. Under these conditions Eulers theorem requires

$$f = x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_n \frac{\partial f}{\partial x_n}.$$

By further restricting u_I to the linear homogeneous subset of all homothetic functions, Euler's theorem can be applied to (3.10), resulting in

$$(3.11) \qquad u_I^1 = u_I^0\left(\mathbf{x_I^0}\right) + \sum_{i \in I} \frac{\partial u_I}{\partial x_i^0} x_i^1 - u_I^0\left(\mathbf{x_I^0}\right).$$

Using the first order conditions of utility maximization [see Equation (2.4)] and simplifying yields

$$(3.12) \qquad u_I^1 = \sum_{i \in I} \lambda p_i^0 x_i^1$$

where λ is the marginal utility of total expenditure. Setting $\lambda=1$ yields a theoretically valid form of the aggregator functions

$$(3.13) \qquad u_I^1\left(\mathbf{x_I^1}\right) = \sum_{i \in I} p_i^0 x_i^1 = \mathbf{p_I^0} \cdot \mathbf{x_I^1}$$

where $\mathbf{p_I^0}$ is the exogenous price vector for group I faced by some base consumer and $\mathbf{x_I^1}$ is the consumption vector from group I for the observed consumer. Following a procedure similar to that outlined in Equations (3.8) – (3.13) results in

$$(3.14) \qquad u_I^0\left(\mathbf{x_I^0}\right) = \sum_{i \in I} p_i^0 x_i^0 = \mathbf{p_I^0} \cdot \mathbf{x_I^0} \,.$$

Diewert (1976) defines an exact quantity aggregator as one that satisfies

(3.15) $$X_I = \frac{u_I^1\left(\mathbf{x_I^1}\right)}{u_I^0\left(\mathbf{x_I^0}\right)}.$$

Substituting (3.13) and (3.14) into (3.15) yields

(3.16) $$X_I = \frac{u_I^1\left(\mathbf{x_I^1}\right)}{u_I^0\left(\mathbf{x_I^0}\right)} = \frac{\mathbf{p_I^0} \cdot \mathbf{x_I^1}}{\mathbf{p_I^0} \cdot \mathbf{x_I^0}}.$$

Then by weak factor reversal, P_I and X_I must satisfy

(3.17) $$P_I \; X_I = \frac{\mathbf{p_I^1} \cdot \mathbf{x_I^1}}{\mathbf{p_I^0} \cdot \mathbf{x_I^0}}.$$

Substituting (3.16) into (3.17) and simplifying yields

(3.18) $$P_I = \frac{\sum_{i \in I} p_i^1 x_i^1}{\sum_{i \in I} p_i^0 x_i^1} = \frac{\mathbf{p_I^1} \cdot \mathbf{x_I^1}}{\mathbf{p_I^0} \cdot \mathbf{x_I^1}}$$

which is the Paasche index. So, not only is the Paasche index appropriate in the absence of complete price data but it also is theoretically justified when the aggregator function u_I is assumed to be a member of the linear homogeneous subset of all homothetic functions. This restriction on u_I is stronger than necessary but does imply that the Paasche index is appropriate when the aggregator function must be homothetic.

When it is desirable for u_i to be non-homothetic, a different price index must be used. A properly specified aggregate demand model requires that the index number used be consistent with a non-homothetic aggregator function. In his important paper on index number theory, Diewert (1976) shows that the TT index is exact for any non-homothetic functional form of the aggregator function (Theorem 2.16, page 122). However, it was shown in Section 3.2 that the TT index results in biased group prices in the absence of complete price data. Given the serious consequences of imposing homotheticity, further research to identify a price index that is both consistent with the assumption of non-homothetic aggregator functions and is unbiased in the absence of complete price data would be a substantial contribution to aggregation theory.

CHAPTER 4: ESTIMATION OF AGGREGATE BEEF, POULTRY, AND PORK DEMANDS

Recently, a considerable amount of research has been directed towards estimating consumer demand for meats [see, for instance, Braschler (1983), Chavas (1983), Wohlgenant (1985), Dahlgran (1987), Moschini and Meilke (1989), Choi and Sosin (1990), Moschini and Vissa (1993), and Eales and Unnevehr (1993)]. Generally, these studies estimate a system of demand equations for beef, poultry, and pork using time-series data that have been aggregated by species and across households. Implicit in each of these models are two separability assumptions.[13] First that the meat group is weakly separable from all other commodities. Second that the individual meat commodities are weakly separable over the beef, poultry, and pork partition. Given these assumptions, the consumer's indirect utility function can be written as

$$(4.1) \qquad V = V\left(P_{bf}, P_{ch}, P_{pk}, m\right)$$

where P_{bf}, P_{ch}, P_{pk} are composite prices for the beef, poultry, and pork groups respectively and m is total expenditure on the three meat varieties. If separability is maintained over the beef, poultry, and pork partition, then estimation of the group demand equations associated with (4.1) requires each composite price to be consistent with the assumption of a homothetic aggregator function (see Section 2.2).[14] However, both separability and homotheticity place severe restrictions on the model and can bias parameter estimates when they are improperly enforced. Therefore, a model that relaxes these assumptions would represent a substantial improvement over previous models.

In Section 4.1, a less restrictive demand model is discussed. Like previous models of meat demand, this model requires the first separability assumption (i.e. that the meat group is weakly separable from all other goods). Unlike previous models, this model does not require the additional assumption that the individual meat commodities

[13] At the time these studies were conducted (i.e. prior to the Generalized Composite Commodity Theorem), the only alternative to separability was to assume that intra-group prices were consistent with the conditions of Hicks – Lenontief aggregation. However, the restrictions imposed by Hicks – Leontief have generally been considered far more restrictive than those imposed by separability and, therefore, have seldom been used in demand research.

[14] Segerson and Mount (1985) developed a model of meat demand that allows for separability but does not require homothetic aggregator functions. However, their approach has not been used in any other meat demand studies.

are weakly separable in the beef, poultry, and pork partition. In addition, the model places no implicit or explicit restrictions on the form of the aggregator functions.

The model proposed in Section 4.1 is only valid when the data being used can be aggregated in a manner consistent with the Generalized Composite Commodity Theorem (GCC). In Section 4.3, a number of data partitions are tested for violations of the GCC. The results give some indication that regional variations in prices may play an important role in creating valid aggregates. The model is used in Section 4.4 to estimate aggregate demand equation parameters for each data partition that is consistent with the conditions of the GCC. The parameter and elasticity estimates derived from this model are contrasted to those derived from the standard model that assumes weak separability of the utility function. Unfortunately, the data partitions that satisfied the conditions of the GCC generally had insufficient degrees of freedom to produce significant parameter estimates. Thus, the results were inconclusive about the effect that enforcing homotheticity has on the demand model.

4.1 The Model

The choice of functional form is an important consideration when modeling demand because the system of demand equations is generally derived from a direct utility, indirect utility, or expenditure function. Unfortunately, theory provides little guidance in choosing a functional form other than to require that the resulting demand equations satisfy the regularity conditions of homogeneity of degree zero in prices and expenditure, adding-up and symmetry. Even these conditions provide little guidance since they can be imposed algebraically on the majority of demand equation systems.

To address the ambiguity involved in choosing a functional form, Diewert (1971) introduced the concept of flexible functional forms (FFF). Diewert (1973d: p. 285) defines a FFF as a function that "contains precisely the number of parameters needed to provide a second order approximation to an arbitrary twice differentiable … function satisfying the appropriate regularity conditions." The advantage to using a FFF is their ability to approximate *any* function within some arbitrarily small neighborhood of the point of approximation. A potential drawback to using FFF's is that the regularity conditions will only hold locally, when all data points fall within this neighborhood

[Caves and Christensen (1980)]. The regularity conditions for non-flexible forms, such as the linear expenditure system (LES), hold globally. Nonetheless, the use of FFF's in empirical demand analysis has grown rapidly since their introduction. In fact, non-flexible forms are seldom found in the modern literature, implying that most researchers find the shortcomings of FFF's relatively minor compared to the advantages.

An example of a FFF found in the demand literature is the Translog form, introduced by Christensen, Jorgenson, and Lau (1975). Consider an arbitrary logarithmic indirect utility function:

$$(4.2) \qquad \ln V = \ln V\left(\frac{\mathbf{p}}{m}\right)$$

where $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ is a vector of group prices, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a corresponding vector of group quantities, and $m = \mathbf{p} \cdot \mathbf{x}$ is total expenditure. A logarithmic second-order Taylor Series approximation to (4.2) yields

$$(4.3) \qquad \ln V = \alpha_0 + \sum_i \alpha_i \ln\frac{p_i}{m} + \tfrac{1}{2}\sum_i \sum_j \beta_{ij} \ln\frac{p_i}{m} \ln\frac{p_j}{m}$$

where α_i and β_{ij} are parameters. Budget share equations, w_i, derived by Roy's identity[15] are of the form

$$(4.4) \qquad w_i = \frac{\alpha_i + \sum_j \beta_{ij} \ln\frac{p_j}{m}}{\sum_j \alpha_j + \sum_j \sum_k \beta_{jk} \ln\frac{p_k}{m}}\ .$$

The model invokes the normalization $\Sigma_j \alpha_j = 1$ although the choice is arbitrary [Christensen, Jorgenson, and Lau (1975)].[16] Homogeneity and adding-up are satisfied by

---

[15] The logarithmic form of Roy's identity is given by $w_i(\mathbf{p}, m) = \dfrac{p_i\, x_i}{m} = -\dfrac{\frac{\partial \ln V(\mathbf{p}, y)}{\partial \ln p_i}}{\frac{\partial \ln V(\mathbf{p}, y)}{\partial \ln m}}$.

[16] The normalization $\Sigma_j \alpha_j = 1$ is convenient because $w_i = \alpha_i$ at the point $(\mathbf{p}, m) = (1, 1)$ [Moschini, Moro, Green (1994)].

the share equations of (4.4) but symmetry requires the additional restriction $\beta_{ij} = \beta_{ji}$ for all *i, j*. Given these restrictions, it is only possible to estimate *(n-1)* of the share equations because all parameters in the final share equation may be derived from the parameter estimates in the other equations.

Two other demand systems that employ flexible forms (and that are commonly used in demand research) are the Rotterdam proposed by Theil (1965) and the Almost Ideal Demand System (AIDS) proposed by Deaton and Muellbauer (1980). Both of these models are similar to the Translog model in that the share equations (demand equations, in the case of the Rotterdam model) are derived from a functional form that satisfies Diewert's definition of flexible. The share equations in the Translog model are derived from a second order approximation to any logarithmic indirect (or direct) *utility* function. In a similar fashion, the AIDS model begins with a second order approximation to any price independent generalized logarithmic (PIGLOG) *expenditure* function and the Rotterdam model begins with a first order approximation to any logarithmic *demand* function.

Although the Translog, AIDS, and Rotterdam models are all based on flexible functional forms, a considerable amount of disagreement exists among researchers about which is the most appropriate for any given demand analysis. Thompson (1988) identifies theoretical and empirical criteria for choosing from among the various FFF's. However, he concludes "whether the benefits of such a formal selection procedure … outweigh the costs in the context of a data set of questionable quality is a judgment that the researcher must make (p. 179)." Given the lack of consensus among researchers, choosing a specific functional form from the class of flexible forms becomes largely a matter of personal preference. However, the Rotterdam model is only applicable when aggregating across households. Since the data in this study were aggregated across commodities, and not households, the Rotterdam model would not be appropriate. In choosing between the translog and AIDS models, no formal selection process was conducted. The translog model was chosen for this study although an AIDS model specification would have been equally appropriate.

When the share equations of (4.4) are estimated for aggregated commodities, it is important to choose group prices, P_I, that are consistent with the underlying assumptions of the model. If commodities are aggregated under the assumption of separable utility, (i.e. the utility function can be written as in (2.9)) then estimation of group demands, X_I, requires the additional restriction of homothetic aggregator functions (see Section 2.2). It was shown in Section 3.3 that the Paasche index is consistent with the assumption of linear homogeneous aggregator functions, which are a subset of the homothetic class of functions. When aggregates are created that satisfy the conditions of the GCC, then estimation of group demands does not require separability. In this case, a non-homothetic aggregator function would be less restrictive than a homothetic aggregator function. As mentioned in Section 3.3, Diewert (1976) has shown that the Törnqvist – Theil (TT) index is consistent with the assumption of non-homothetic aggregator functions.[17]

4.2 Data

The National Livestock and Meat Board provided monthly consumption and expenditure data for 1057 consumers during the period from November 1993 to March 1994 for twenty-six varieties of meat. The various meat varieties can be broadly identified as beef, poultry, or pork products as shown in Table 4.1. To simplify estimation procedures, it was necessary to use only those observations that consumed from all three of the meat varieties (i.e. beef, poultry, and pork) during each of the five months.[18] However, this procedure required discarding a large number of households. To avoid this problem, and increase the number of usable observations, the consumption and expenditure data were summed over the five months and treated as a cross-section. This reduced the number of households dropped from the sample because households were only required to consume from each meat group during the entire five-month period.

[17] However, it is important to recall that the TT index produces biased group prices in the absence of complete price data (see Section 3.2).
[18] Note that this would not be necessary in the context of a limited dependent variable model.

Table 4.1: Composition of Beef, Poultry, and Pork Groups

Beef	Poultry	Pork
Ground Beef *(bf_1)*	Whole Chicken *(ch_1)*	Roasts *(pk_1)*
Premium Roast *(bf_2)*	Chicken Breasts *(ch_2)*	Chops *(pk_2)*
Non-premium Roast *(bf_3)*	Ground Chicken *(ch_3)*	Hams *(pk_3)*
Other Roast *(bf_4)*	Other Chicken *(ch_4)*	Ribs *(pk_4)*
Premium Steak *(bf_5)*	Whole Turkey *(ch_5)*	Sausage *(pk_5)*
Non-premium Steak *(bf_6)*	Turkey Breasts *(ch_6)*	Steaks / Cutlets *(pk_6)*
Other Steak *(bf_7)*	Ground Turkey *(ch_7)*	Other Pork *(pk_7)*
Stew Meat *(bf_8)*	Other Turkey *(ch_8)*	
Other Beef *(bf_9)*	Cornish Hens *(ch_9)*	
	Other Poultry *(ch_0)*	

Note: Each column groups individual meat commodities by species. The data consisted of nine beef varieties, ten poultry varieties, and seven pork varieties. The values in parentheses are shorthand for referring to the various meat varieties (i.e. bf_1 refers to ground beef).

Since it is unlikely that any preference change occurred during this short period of data collection, this summation is appropriate.

Unfortunately, many surveys had incomplete price information. There were 757 observations that reported purchases of one or more of the meat varieties at zero expenditure. These observations were deleted because the imputation of commodity prices requires both data elements. After completing all the necessary deletions, only155 usable observations remained.

4.3 Verifying that the Aggregates are Consistent with the Conditions of GCC

Before estimating (4.4), it is necessary to validate the aggregation of the individual meat commodities into the broad beef, poultry, and pork groups.[19] Recall that the GCC requires independence of the random variables ρ_i and R_I for all i where

$$\rho_{i\in I} = \log\left(\frac{p_i}{P_I}\right),$$

$R_I = log\ (P_I)$, and P_I represents group price as determined by an appropriate price index.[20] It is well known that if ρ_i and R_I are independent random variables, then

$$correlation(\rho_{i\in I}, R_I) = 0.$$

However, the converse of this statement is not generally true (i.e. *correlation* $(\rho_i, R_I) = 0$ *does not* imply that ρ_i and R_I are independent random variables). Kapur and Kesavan (1992) show that *correlation* $(\rho_i, R_I) = 0$ is only a necessary and sufficient condition for independence when ρ_i and R_I are distributed multivariate normal. When the normality condition is not met, *correlation* $(\rho_i, R_I) = 0$ is still necessary for independence but no longer sufficient. No a priori assumptions about the distribution of the data are made in this study and, therefore, the tests of independence conducted in this section only provide a necessary condition for creating aggregates consistent with the GCC.[21]

[19] Note that other potential, and economically meaningful, groupings exist but the beef, poultry, and pork partition is ubiquitous in the literature.
[20] The Törnqvist – Theil index was used in the following tests.
[21] Recent research on entropy may provide a necessary and sufficient test for independence that does not rely upon normality.

The correlation coefficients for the 26 meat varieties and their respective groups are given in Table 4.2.[22] The numbers in parentheses represent the probability that *correlation* (ρ_i, R_I) = 0. Define this probability as Ω. A value of $\Omega = 0.05$ implies that *correlation* (ρ_i, R_I) is significantly *different* from zero at the 95% confidence level. Thus, any value of $\Omega \leq 0.05$ implies rejection of the test for independence at the 95% confidence level. However, for multiple hypothesis tests to be jointly significant at the 95% confidence level, an adjusted value of Ω is required. A common adjustment, originally proposed by Sidak, is given by

$$\Omega \leq 1 - (1 - \alpha)^{\frac{1}{n}} \tag{4.5}$$

where α is the desired confidence level and n is the number of hypothesis being jointly tested. As shown in Table 4.1, the beef group consists of nine beef varieties, the poultry group consists of ten poultry varieties and the pork group consists of seven pork varieties. Substituting these values of n into (4.5) and setting $\alpha = 0.95$ yields the following values of Ω required for *correlation* (ρ_i, R_I) to be jointly significantly *different* from zero at the 95% confidence level:

$$\Omega_{\mathrm{bf}} \leq 1 - (1 - 0.95)^{\frac{1}{9}} \Rightarrow \Omega_{\mathrm{bf}} \leq 0.0057$$
$$\Omega_{\mathrm{ch}} \leq 1 - (1 - 0.95)^{\frac{1}{10}} \Rightarrow \Omega_{\mathrm{ch}} \leq 0.0051$$
$$\Omega_{\mathrm{pk}} \leq 1 - (1 - 0.95)^{\frac{1}{7}} \Rightarrow \Omega_{\mathrm{pk}} \leq 0.0073$$

where Ω_{bf}, Ω_{ch}, Ω_{pk}, are the required values of Ω for the beef, poultry, and pork groups respectively. Thus, *correlation* (ρ_i, R_I) is significantly different from zero at the 95%

[22] The correlation coefficients listed in Table 4.2 were adjusted to account for potential variations in regional mean prices. See Appendix 2 – SAS Code for more detail on the method used to calculate the correlation coefficients.

Table 4.2: Correlation Coefficients for the Beef, Poultry, and Pork Partition

	R_{bf}
ρ_{bf1}	.2435 (.0037)
ρ_{bf2}	.2130 (.0724)
ρ_{bf3}	.0845 (.4772)
ρ_{bf4}	.1880 (.5798)
ρ_{bf5}	-.1514 (.2010)
ρ_{bf6}	-.1043 (.3903)
ρ_{bf7}	-.2558 (.3989)
ρ_{bf8}	-.6169 (.0001)
ρ_{bf9}	.2555 (.0399)

	R_{ch}
ρ_{ch1}	-.2922 (.0099)
ρ_{ch2}	.5444 (.0001)
ρ_{ch3}	.9136 (.0040)
ρ_{ch4}	.1756 (.1124)
ρ_{ch5}	.1686 (.1903)
ρ_{ch6}	.4422 (.0185)
ρ_{ch7}	-.1018 (.5924)
ρ_{ch8}	-.5132 (.0123)
ρ_{ch9}	-.8084 (.0084)
ρ_{ch0}	.1219 (.5704)

	R_{pk}
ρ_{pk1}	.1424 (.2655)
ρ_{pk2}	.2230 (.0281)
ρ_{pk3}	.4161 (.0001)
ρ_{pk4}	-.4469 (.0010)
ρ_{pk5}	.-.3162 (.0211)
ρ_{pk6}	.1658 (.4085)
ρ_{pk7}	-.0179 (.9038)

Note: The values in this table represent the correlation coefficient between the two variables. The numbers in parenthesis represent the probabilities that the correlation is equal to zero. The subscripts refer to the individual meat commodities defined in Table 4.1. Group prices were calculated using the Törnqvist – Theil price index. There were 155 observations.

confidence level for ρ_{bf1} and R_{bf}, ρ_{bf8} and R_{bf}, ρ_{ch2} and R_{ch}, ρ_{ch3} and R_{ch}, ρ_{pk3} and R_{pk}, and ρ_{pk4} and R_{pk}. These results indicate dependence for each of these pairs of random variables. Thus, aggregation of commodities into beef, poultry, and pork groups is not valid under the conditions of the GCC for the complete data set.

Although partitioning by species (i.e. beef, poultry, and pork) has intuitive appeal, the choice of groups is arbitrary and a number of other economically meaningful partitions exist. Table 4.3 lists three alternative partitions.[23] Partition I groups commodities by species as described above. Partition II groups the commodities by their fat content based upon the consumers *perceptions* about the fat content of the various meats. Meats perceived to contain less than 16 percent fat were put into the low fat group. Those perceived to contain 16 to 25 percent fat were put into the medium fat group. Those perceived to contain greater than 25 percent fat were put into the high fat group. In partition III, the commodities are grouped by their price based upon the survey data. The mean meat price for all varieties was $1.60 per pound. Meats priced below the mean were put into the low priced group while meats priced above the mean were put into the high priced group. The correlation coefficients were calculated for each set of variables for each of the partitions. The violations of the GCC from each partition are summarized in the column labeled *all* in Table 4.4, showing that the conditions of the GCC were not satisfied for any of the partitions over the entire data set. One potential reason that the proposed data groupings failed to satisfy the conditions of the GCC is a failure to account for regional price differences in the data. Nicol (1991) and Diewert (1995) argue that regional effects can impact commodity prices enough to justify treating identical commodities in different regions as unique products. To allow for spatial differences, the data were partitioned into the following regions: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.[24] The correlation coefficients were

[23] The number of potential groupings of 26 commodities is vast and these three partitions only represent a small subset of the total. Lewbel (1996) suggested using cluster analysis to identify appropriate groups. However, this approach is only useful when the identified groups are economically meaningful. Davis (1998) used cluster analysis in an attempt to identify data partitions consistent with the GCC. Not only did the groups identified have little economic meaning, they also did not satisfy the conditions of the GCC.
[24] The National Livestock and Meat Board did not provide specific definitions of these regions.

Table 4.3: Various Groupings of the Meat Varieties

	Partition I (Species)	Partition II (Fat)	Partition III (Price)
BEEF PRODUCTS			
Ground Beef	A	C	A
Premium Roast	A	B	A
Non-premium Roast	A	B	A
Other Roast	A	B	A
Premium Steak	A	B	A
Non-premium Steak	A	B	A
Other Steak	A	B	A
Stew Meat	A	B	A
Other Beef	A	B	A
POULTRY PRODUCTS			
Whole Chicken	B	A	B
Chicken Breasts	B	A	A
Ground Chicken	B	A	A
Other Chicken	B	A	B
Whole Turkey	B	A	B
Turkey Breasts	B	A	B
Ground Turkey	B	A	B
Other Turkey	B	A	B
Cornish Hens	B	A	A
Other Poultry	B	A	B
PORK PRODUCTS			
Roasts	C	C	A
Chops	C	B	A
Hams	C	C	A
Ribs	C	C	A
Sausage	C	C	A
Steaks / Cutlets	C	C	B
Other Pork	C	C	A
TOTAL GROUPS	3	3	2

Note: For each partition, commodities with the same letter comprise a group. Explanations of the Groups are given in Table 4.4.

Table 4.4: Summary of Tests for Violations of the GCC

PARTITION I	Number of Goods	Violations of the GCC by Region									
		All	NE	MA	ENC	WNC	SA	ESC	WSC	M	P
Beef (A)	9	2	0	0	0	0	2	0	0	0	0
Poultry (B)	10	2	0	0	0	0	1	1	0	0	1
Pork (C)	7	2	0	0	0	0	1	1	0	0	1
Observations	--	155	15	16	27	8	24	10	14	17	24
PARTITION II											
Low Fat Meats (A)	10	1	0	0	0	0	1	0	0	0	1
Medium Fat Meats (B)	9	4	0	2	1	0	0	0	0	1	0
High Fat Meats (C)	7	6	4	2	5	1	6	3	3	3	4
Observations	--	170	13	22	29	8	27	13	16	15	27
PARTITION III											
High Price Meats (A)	18	16	7	7	12	1	12	6	7	8	11
Low Price Meats (B)	8	1	0	0	0	0	1	0	0	0	0
Observations	--	203	16	33	33	11	32	15	13	20	30

Note: All = All States, NE = New England States, MA = Mid Atlantic States, ENC = East North Central States, WNC = West North Central States, SA = South Atlantic States, ESC = East South Central States, WSC = West South Central States, M = Mountain States, and P = Pacific States. Differences in the number of observations between partitions are due to differences in how the groups are defined. Recall that observations not consuming at least one good from each group were deleted.

calculated for each partition over each of the regions. Table 4.4 summarizes the violations of the GCC that were found. Zero violations were found in six regions using partition I. These subsets of the data satisfy a *necessary* condition for consistency with the GCC. It is important to note that zero violations is not a *sufficient* condition for creating aggregates under the GCC because it is possible for two random variables to have a zero correlation coefficient and still be dependent.

4.4 Results

To establish if imposing homotheticity on the model impacts parameter and elasticity estimates, the share equations defined by (4.4) were estimated using Partition I for each of the six regions with zero violations. Parameter estimates were derived using theiterative seemingly unrelated regression (ITSUR) procedure. The models were first estimated with group prices, P_I, defined using the Törnqvist – Theil (TT) price index, which is consistent with a non-homothetic preference structure for intra-group commodities. Next, the models were estimated with group prices defined using the Paasche price index, which was shown to be consistent with a linear homogeneous preference structure. This is a stronger condition than necessary but does impose homotheticity on intra-group preferences. Unfortunately, the model could not be estimated in three of the regions and the parameter estimates were generally found to have little significance in the remaining regions. The West North Central region was not of full rank implying there were insufficient degrees of freedom to estimate the parameters of the model. In the New England and East North Central regions, the model failed to converge. This could indicate that the share equations do not fit the data or could be attributable to the limited degrees of freedom in each region. The model did converge for the remaining three regions: the Mid Atlantic, Mountain, and West South Central regions. However, there were only a small number of observations from each of these regions (see Table 4.4) implying that the robustness of these estimates is questionable. The parameter estimates from both the linear homogeneous (i.e. Paasche index) model and non-homothetic (i.e. TT index) model are listed in Table 4.5 with t-values in parentheses for each region. An examination of the t-values indicates that the partition with the most significant parameter estimates come from the Mid Atlantic

Table 4.5: Parameter Estimates for the Regions that Satisfy the Conditions of the GCC

	Mid Atlantic		***Mountain***		***West South Central***	
Param.	***TT Index***	***Paasche Index***	***TT Index***	***Paasche Index***	***TT Index***	***Paasche Index***
$\alpha 1$	0.63 (1.99)	0.46 (7.43)	0.24 (0.22)	0.19 (0.17)	-0.02 (-0.03)	-0.05 (-0.06)
$\alpha 2$	0.07 (-0.14)	0.32 (5.30)	0.44 (0.60)	0.51 (0.61)	0.61 (1.31)	0.77 (0.94)
$\beta 11$	0.48 (0.82)	0.08 (1.33)	-0.10 (-0.12)	-0.27 (-0.23)	0.16 (0.59)	0.04 (0.15)
$\beta 12$	-0.41 (-0.75)	0.007 (0.15)	-0.27 (-0.24)	-0.03 (-0.10)	-0.09 (-0.48)	-0.08 (-0.41)
$\beta 13$	0.01 (0.08)	0.009 (0.53)	0.21 (0.29)	0.15 (0.34)	-0.17 (-0.51)	-0.10 (-0.29)
$\beta 22$	0.31 (0.91)	0.03 (0.78)	0.60 (0.31)	0.28 (0.36)	0.10 (0.60)	0.22 (0.65)
$\beta 23$	-0.02 (-0.06)	0.03 (2.10)	-0.32 (-0.30)	-0.22 (-0.32)	0.09 (0.56)	-0.002 (-0.01)
$\beta 33$	0.07 (0.52)	0.009 (1.40)	0.10 (0.33)	0.08 (0.38)	0.12 (0.36)	0.09 (0.28)

Note: There were 16 observations from the Mid Atlantic region, 17 observations from the Mountain region, and 14 observations from the West South Central Region. Parameter estimates were generated using ITSUR. The numbers in parentheses represent t-values.

region with the Paasche index. None of the regions produced significant parameter estimates using the TT index.

Price and expenditure elasticities were calculated for the Mid Atlantic region since the parameter estimates from this region generally had the highest significance level.[25] These estimates are summarized in Table 4.6 for the linear homogeneous model and in Table 4.7 for the non-homothetic model. The elasticity estimates from the linear homogeneous model are similar to those found in previous studies [see Moschini and Vissa (1994)] but the elasticity estimates from the non-homothetic model are radically different. It may be tempting to attribute these differences solely to the additional preference restrictions placed upon the linear homogeneous model but there are many reasons not to do so.

First, the elasticity estimates are highly dependent upon the parameter estimates. Therefore, the low significance of the parameter estimates from the non-homothetic model makes the resulting elasticity estimates highly questionable. Second, as mentioned in Chapter 3, the TT index is biased in the absence of complete price data. Depending on the amount of bias, the parameter and elasticity estimates from the non-homothetic model may have been significantly impacted. Third, any individuals that did not consume from all three meat categories were deleted. These deletions were necessary given the model used but would not be required if a limited dependent variable model were specified.

---

[25] Own price (ξ_{ii}), cross price (ξ_{ij}), and expenditure (η_i) elasticities corresponding to (4.4) are given by:

$$\xi_{ii} = -1 + \frac{\beta_{ii}}{\alpha_i + \sum_j \beta_{ij} \ln \frac{p_j}{m}} - \frac{\sum_j \beta_{ij}}{\sum_j \alpha_j + \sum_j \sum_k \beta_{jk} \ln \frac{p_k}{m}}$$

$$\xi_{ij} = \frac{\beta_{ij}}{\alpha_i + \sum_j \beta_{ij} \ln \frac{p_j}{m}} - \frac{\sum_j \beta_{ij}}{\sum_j \alpha_j + \sum_j \sum_k \beta_{jk} \ln \frac{p_k}{m}}$$

$$\eta_i = 1 - \frac{\sum_j \beta_{ij}}{\alpha_i + \sum_j \beta_{ij} \ln \frac{p_j}{m}} + \frac{\sum_j \sum_k \beta_{jk}}{\sum_j \alpha_j + \sum_j \sum_k \beta_{jk} \ln \frac{p_k}{m}}.$$

Table 4.6: Price and Expenditure Elasticities for the Mid Atlantic Region Using the Paasche Price Index

	Beef	***Poultry***	***Pork***	***Expenditure***
Beef	-0.79 (-2.18)	-0.12 (-0.48)	-0.07 (.85)	0.98 (47.55)
Poultry	-0.19 (-0.55)	-0.94 (-3.68)	0.10 (1.07)	1.03 (40.06)
Pork	-0.14 (-0.73)	0.15 (1.18)	-1.01 (-14.04)	1.00 (56.55)

Note: There were 16 observations from the Mid Atlantic region. The numbers in parentheses represent t-values.

Table 4.7: Price and Income Elasticities for the Mid Atlantic Region Using the Törnqvist – Theil Price Index

	Beef	***Poultry***	***Pork***	***Expenditure***
Beef	0.37 (0.30)	-.32 (-.81)	-0.05 (-0.10)	.81 (5.38)
Poultry	-1.58 (-1.79)	0.26 (.39)	-0.12 (-0.30)	1.44 (4.95)
Pork	-0.05 (-0.06)	0.09 (0.17)	-0.76 (-1.35)	0.71 (3.45)

Note: There were 16 observations from the Mid Atlantic region. The numbers in parentheses represent t-values.

Fourth, the model was estimated with very limited degrees of freedom. Only sixteen observations were available for estimation of the eight parameters. Finally, the tests that were conducted to determine if aggregation under the GCC was valid are incomplete. The condition, *correlation* $(\rho_i, R_I) = 0$ is necessary but *not* sufficient for establishing the independence of ρ_i and R_I. Therefore, it is quite possible that the conditions of the GCC do not actually hold for the Mid Atlantic region (or any of the other regions with zero violations). Because of these many difficulties, it is not possible to make any firm conclusions about the effects that restricting intra-group preferences to linear homogeneous forms had on the results.

CHAPTER 5: CONCLUSIONS

The aggregation of commodities into a smaller number of seemingly related groups is typically necessary to make models of consumer demand empirically tractable. This study has presented both necessary and sufficient conditions for creating theoretically valid aggregates using three alternative methods. In addition, price indices consistent with each of these aggregation methods have been suggested. This research is important because the elasticity estimates resulting from aggregate demand models provide market analysts with objective information about consumer purchasing behavior. Inappropriate aggregation can lead to biased elasticity estimates and ineffective policies.

The most common approach to the aggregation problem assumes weak separability of the utility function for some data partition. This is equivalent to saying that consumer behavior is consistent with a two-stage budgeting process where income is first distributed to broad commodity groups and then to the individual commodities comprising those groups. Under weak separability of the utility function, consumer demands for *individual* commodities can be expressed as functions of intra-group commodity prices and group expenditure. With the additional assumption that consumer preferences for intra-group commodities are homothetic, consumer demands for *composite* commodities can be expressed as functions of group prices and total expenditure. This approach to modeling demand is quite restrictive because homotheticity requires intra-group budget shares to be independent of group expenditure and separability severely limits the potential substitution effects between inter-group commodities.

The Hicks – Leontief Composite Commodity Theorem rationalizes the existence of group demand equations without separability or homotheticity but requires that relative intra-group commodity prices (ρ_i) remain constant across time (and/or consumers). This is equivalent to saying that intra-group commodity prices are perfectly correlated (i.e. $correlation(\rho_i, \rho_j) = 1 \;\forall i, j$) and is generally considered far more restrictive than weak separability and homotheticity. Lewbel (1996) proposed a generalized composite commodity theorem (GCC) that relaxes the conditions of Hicks – Leontief aggregation by allowing the ρ_i*'s* to vary and only requires independence between each ρ_i and its corresponding group price (R_I). Lewbel found this independence condition easy to satisfy when using highly aggregated time-series data. However, this study confirms earlier results by Davis (1998) that satisfying the conditions of the GCC can be difficult at lower levels of aggregation.

Another important consideration when estimating aggregate demand models is the form of the price index. The specified price index must be consistent with the assumption made about the structure of intra-group preferences (i.e. homothetic or non-homothetic) and must be appropriate for the data being used. An important finding of this study is that many commonly employed price indices (i.e. the Ideal, Laspeyres, and Törnqvist – Theil) are biased in the absence of complete price data. Only the Paasche index was found unbiased when unobserved prices were set to arbitrary values. Furthermore, it was shown that the Paasche index is consistent with linear homogeneous intra-group preferences, which are a subset of the broader class of homothetic preferences, and is therefore appropriate when aggregates are created under the assumption of weak separability. Diewert (1976) has shown that the Törnqvist – Theil (TT) index is consistent with non-homothetic intra-group preferences. Therefore, when aggregates satisfy the conditions of the GCC the TT index may yield a more general model of aggregate demand by allowing for non-linear Engel curves. However, it is important to recall that the TT price index is biased in the absence of complete price data.

Using cross-sectional data, individual meat varieties were aggregated over three partitions. First by species (i.e. beef, poultry, and pork), second by fat content (i.e. high fat meats, medium fat meats, and low fat meats), and finally by price (i.e. high price meats and low fat meats). Violations of the GCC were found in each partition. The data were then subdivided into nine geographic regions and each partition was again tested for violations of the GCC by region. Violations of the GCC were found in every region under the fat content and price partitions. However, aggregation by species was found consistent with the GCC in six of the nine geographic regions. This implies that regional price differences may be an important consideration when creating aggregates that are consistent with the GCC.

A system of demand equations was estimated for each of the data partitions where there were zero violations of the GCC. The system was first estimated with intra-group preferences restricted to the linear homogeneous subset of the broader class of homothetic preferences. Next, the system was estimated without placing any implicit or explicit restrictions on intra-group preferences. Unfortunately, significant parameter estimates could only be derived from the linear homogeneous model for one of the regions and the parameter estimates from the non-homothetic model were not significant in any region. Therefore, it was not possible to determine if restricting intra-group preferences to linear homogeneous forms had any impact on the price and expenditure elasticities for this data.

This study provides four directions for future research. First, further research is needed to determine which factors most influence the independence of ρ_i and R_I. Some potential factors include the amount of correlation between ρ_i and ρ_j for $i \neq j$, the relative size of the commodity budget shares, and whether the intra-group commodities are substitutes, complements, or a combination. Second, the use of cross-sectional data with incomplete price information is becoming more common in applied demand analysis. The Paasche index is unbiased in the absence of complete price data but is only appropriate when intra-group preferences are linear homogeneous. Therefore, further research to determine a price index that is both unbiased by incomplete price data and consistent with the assumption of non-homothetic preferences would be a substantial contribution to the field of aggregation theory. Third, since the results of this study were inconclusive, further research is required to determine if restricting intra-group preferences to homothetic forms has a significant impact on elasticity estimates. A final suggestion for future research is to determine a necessary and sufficient test for independence of ρ_i and R_I that does not rely upon normality. When ρ_i and R_I are *not* distributed multivariate normal, the test proposed in this study, $correlation(\rho_i, R_I) = 0$, is only a necessary condition for independence. A potential solution to this problem may reside in the field of entropy theory.

REFERENCES

Blackorby, C., Primont, D., and Robert Russell. *Duality, Separability, and Functional Structure: Theory and Economic Applications.* New York: Elsevier/North-Holland, 1978.

Braschler, C. "The Changing Demand Structure for Pork and Beef in the 1970s: Implications for the 1980s." *Southern Journal of Agricultural Economics* 15(1983): 105 – 110.

Alston, J.M., and J.A. Chalfant. "Can We Take the Con Out of Meat Demand Studies?" *Western Journal of Agricultural Economics* 16(1): 36 – 48.

Caves, D.W., and L.R. Christensen. "Global Properties of Flexible Functional Forms." *Am. Econ. Rev.* 70 (3): 422 – 432.

Chavas, J.-P. "Structural Change in the Demand for Meat." *American Journal of Agricultural Economics* 65(1983): 148 – 153.

Choi, S., and K. Sosin. "Testing for Structural Change: The Demand for Meat." *American Journal of Agricultural Economics* 72(1990): 227 – 236.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau. "Transcendental Logarithmic Utility Functions." *Am. Econ. Rev.* 65(3): 367 – 383.

Dahlgran, R.A. "Complete Flexibility Systems and the Stationarity of U.S. Meat Demands." *Western Journal of Agricultural Economics* 12(1987): 152 – 163.

Davis, G.C. "Testing Aggregation Without Separability in Meat Demand: An Investigation of the Generalized Composite Commodity Theorem." Paper presented at the annual American Agricultural Economics Association Meeting, August 1998.

Deaton, A., and John Muellbauer. *Economics and Consumer Behavior.* Cambridge: Cambridge University Press, 1980.

Diewert, W.E. "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function." *Journal of Political Economy* 79(1971): 481 – 507.

Diewert, W.E. "Functional Forms for Profit and Transformation Functions." *Journal of Economic Theory* 6(1973): 284 – 316.

Diewert, W.E. "Exact and Superlative Index Numbers." *J. Econometrics* 4(1976): 115 – 145.

Diewert, W.E. "Index Numbers." *New Palgrave: A Dictionary of Economics*, Vol. 2, J. Eatwell, M. Milgate and P. Newman (eds.), The Macmillan Press, 1987, pp. 767 – 780.

Diewert, W.E. and T.J. Wales. "Flexible Functional Forms and Tests of Homogeneous Separability." *J. Econometrics* 67(1995): 259 – 302.

Eales, J.S. and L.J. Unnevehr. "Simultaneity and Structural Change in U.S. Meat Demand." *American Journal of Agricultural Economics* 75(1993): 259 – 268.

Fischer, I. *The Making of Index Numbers.* Houghton Mifflin, Boston, Mass., 1922.

Frisch, R. "Annual Survey of General Economic Theory: The Problem of Index Numbers," *Econometrica* 4(1936): 1 – 39.

Gorman, W.M. "Separable Utility and Aggregation." *Econometrica* 27(1959): 469 – 481.

Guilkey, D.K., et al. "A Comparison of the Performance of Three Flexible Functional Forms." *International Economic Review* 24(3): 591 – 616.

Hicks, J.R. *Value and Capital.* Oxford: Oxford University Press, 1936.

Hicks, J.R. "The Valuation of the Social Income," *Economica 7*(1940): 105 – 124.

Kapur, J.N. and H.K. Kesavan. *Entropy Optimization Principles with Applications*. Boston: Academic Press, Inc., 1992.

Keynes, J.M. *Treatise on Money, Vol. 1*, London: Macmillan.

Leontief, W. "Composite Commodities and the Problem of Index Numbers." *Econometrica, 4* (1936): 39 – 59.

Lewbel, A. "Aggregation without Separability: A Generalized Composite Commodity Theorem." *Am. Econ. Rev.* 96 (3): 524 – 543.

Moschini, G. and K.D. Meilke. "Modeling the Pattern of Structural Change in U.S. Meat Demand." *American Journal of Agricultural Economics* 71 (1989): 253 – 261.

Moschini, G. and Anuradha Vissa. "Flexible Specification of Mixed Demand Systems." *American Journal of Agricultural Economics* 75 (1993): 1 – 9.

Moschini, G., D. Moro, and R.D. Green. "Maintaining and Testing Separability in Demand Systems." *American Journal of Agricultural Economics* 76(1994): 61 – 73.

Nicol, C. J. "The Effect of Expenditure Aggregation on Hypothesis Tests in Consumer Demand Systems." *International Economic Review* 32 (2): 405 – 416.

Segerson, K. and Timothy D. Mount. "A Non-Homothetic Two-Stage Decision Model Using AIDS." *Review of Economics and Statistics* 67 (4): 630 – 639.

Theil, H. “The Information Approach to Demand Analysis.” *Econometrica*, 33(1965): 67 – 87.

Thompson, G.D. “Choice of Flexible Functional Forms: Review and Appraisal.” *Western Journal of Agricultural Economics* 13 (2): 169 – 183.

Wohlgenant, M.K. “Estimating Cross Elasticities of Demand for Beef.” *Western Journal of Agricultural Economics* 10 (2): 322 – 329.

APPENDIX A: PROOFS OF AXIOMS

In this appendix, the Ideal, Paasche, Laspeyres, and Törnqvist – Theil indices are tested against the nine axioms suggested in Chapter 3. Diewert (1987) identifies the axioms that are violated by each of these indices. However, he does not explicitly prove his results. The proofs offered in this appendix serve to verify the findings of Diewert. Indices failing an axiom are highlighted in bold.

<u>Axiom 1 (Identity):</u> Assume $\alpha > 0$, $\beta > 0$, $\mathbf{p^0} = \mathbf{p^1}$ and $\mathbf{x^0} = \mathbf{x^1}$. Then

$$P\left(\mathbf{p^0}, \mathbf{p^1}, \alpha\mathbf{x^0}, \beta\mathbf{x^1}\right) = 1.$$

1. Ideal Index

$$\mathbf{p^1} = \mathbf{p^0} \Rightarrow \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^0}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^0}\right)} = 1 \text{ and } \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^1}\right)} = 1 \Rightarrow \sqrt{\frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^0}\right)\left(\mathbf{p_I^1} \cdot \mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^0}\right)\left(\mathbf{p_I^0} \cdot \mathbf{x_I^1}\right)}} = 1.$$

2. Paasche Index

$$\mathbf{p^1} = \mathbf{p^0} \Rightarrow \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^1}\right)} = 1.$$

3. Laspeyres Index

$$\mathbf{p^1} = \mathbf{p^0} \Rightarrow \frac{\left(\mathbf{p_I^1} \cdot \mathbf{x_I^0}\right)}{\left(\mathbf{p_I^0} \cdot \mathbf{x_I^0}\right)} = 1.$$

4. Törnqvist – Theil Index

$$\mathbf{p^1} = \mathbf{p^0} \Rightarrow \left(\frac{p_i^1}{p_i^0}\right) = 1 \forall i \Rightarrow \prod_{i \in I} \left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{p_i^1 x_i^1}{\mathbf{p^1} \cdot \mathbf{x^1}} + \frac{p_i^0 x_i^0}{\mathbf{p^0} \cdot \mathbf{x^0}}\right)} = 1.$$

<u>Axiom 2 (Proportionality):</u> Assume $\alpha > 0$. *Then*

$$P\left(\mathbf{p^0}, \alpha\mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right) = \alpha P\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right).$$

1. Ideal Index

$$P^{id}\left(\mathbf{p^0}, \alpha\mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right) = \sqrt{\frac{\left(\alpha\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)\left(\alpha\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}} = \sqrt{\frac{\alpha^2\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}}$$
$$= \alpha\sqrt{\frac{\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}} = \alpha P^{id}\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right).$$

2. Paasche Index

$$P^{pa}\left(\mathbf{p^0}, \alpha\mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right) = \frac{\left(\alpha\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)} = \frac{\alpha\left(\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)} = \alpha P^{pa}\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right).$$

3. Laspeyres Index

$$P^{la}\left(\mathbf{p^0}, \alpha\mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right) = \frac{\left(\alpha\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)} = \frac{\alpha\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)} = \alpha P^{la}\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right).$$

4. Törnqvist – Theil Index

$$P^{tt}\left(\mathbf{p^0}, \alpha\mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right) = \prod_{i\in I}\left(\frac{\alpha p_i^1}{p_i^0}\right)^{.5\left(\frac{\alpha p_i^1 x_i^1}{\alpha\mathbf{p^1}\cdot\mathbf{x^1}} + \frac{p_i^0 x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}\right)}$$
$$= \alpha\prod_{i\in I}\left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{p_i^1 x_i^1}{\mathbf{p^1}\cdot\mathbf{x^1}} + \frac{p_i^0 x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}\right)} = \alpha P^{tt}\left(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}\right).$$

Axiom 3 (Invariance to Changes in Scale): Assume $\alpha > 0$, $\beta > 0$, and $\gamma > 0$. Then

$$P(\alpha\mathbf{p^0}, \alpha\mathbf{p^1}, \beta\mathbf{x^0}, \gamma\mathbf{x^1}) = P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

1. Ideal Index

$$P^{id}(\alpha\mathbf{p^0}, \alpha\mathbf{p^1}, \beta\mathbf{x^0}, \gamma\mathbf{x^1}) = \sqrt{\frac{(\alpha\mathbf{p_I^1}\cdot\beta\mathbf{x_I^0})(\alpha\mathbf{p_I^1}\cdot\gamma\mathbf{x_I^1})}{(\alpha\mathbf{p_I^0}\cdot\beta\mathbf{x_I^0})(\alpha\mathbf{p_I^0}\cdot\gamma\mathbf{x_I^1})}} = \sqrt{\frac{\alpha\beta(\mathbf{p_I^1}\cdot\mathbf{x_I^0})\alpha\gamma(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{\alpha\beta(\mathbf{p_I^0}\cdot\mathbf{x_I^0})\alpha\gamma(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}}$$

$$= \sqrt{\frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}} = P^{id}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

2. Paasche Index

$$P^{pa}(\alpha\mathbf{p^0}, \alpha\mathbf{p^1}, \beta\mathbf{x^0}, \gamma\mathbf{x^1}) = \frac{(\alpha\mathbf{p_I^1}\cdot\gamma\mathbf{x_I^1})}{(\alpha\mathbf{p_I^0}\cdot\gamma\mathbf{x_I^1})} = \frac{\alpha\gamma(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{\alpha\gamma(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} = \frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} = P^{pa}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

3. Laspeyres Index

$$P^{la}(\alpha\mathbf{p^0}, \alpha\mathbf{p^1}, \beta\mathbf{x^0}, \gamma\mathbf{x^1}) = \frac{(\alpha\mathbf{p_I^1}\cdot\beta\mathbf{x_I^0})}{(\alpha\mathbf{p_I^0}\cdot\beta\mathbf{x_I^0})} = \frac{\alpha\beta(\mathbf{p_I^1}\cdot\mathbf{x_I^0})}{\alpha\beta(\mathbf{p_I^0}\cdot\mathbf{x_I^0})} = \frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})} = P^{la}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

4. Törnqvist – Theil Index

$$P^{tt}(\alpha\mathbf{p^0}, \alpha\mathbf{p^1}, \beta\mathbf{x^0}, \gamma\mathbf{x^1}) = \prod_{i\in I}\left(\frac{\alpha p_i^1}{\alpha p_i^0}\right)^{.5\left(\frac{\alpha p_i^1 \gamma x_i^1}{\alpha\mathbf{p^1}\cdot\gamma\mathbf{x^1}} + \frac{\alpha p_i^0 \beta x_i^0}{\alpha\mathbf{p^0}\cdot\beta\mathbf{x^0}}\right)}$$

$$= \prod_{i\in I}\left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{\alpha\gamma p_i^1 x_i^1}{\alpha\gamma(\mathbf{p^1}\cdot\mathbf{x^1})} + \frac{\alpha\beta p_i^0 x_i^0}{\alpha\beta(\mathbf{p^0}\cdot\mathbf{x^0})}\right)}$$

$$= \prod_{i\in I}\left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{p_i^1 x_i^1}{(\mathbf{p^1}\cdot\mathbf{x^1})} + \frac{p_i^0 x_i^0}{(\mathbf{p^0}\cdot\mathbf{x^0})}\right)} = P^{tt}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

Axiom 4 (Invariance to Changes in Units): Assume $\mathbf{p^i}$ and $\mathbf{x^i}$ are n-vectors for i = 0, 1. Define

$$\tilde{\alpha} \equiv (\alpha_1, \alpha_2, \ldots, \alpha_n), \text{ and}$$
$$\tilde{\alpha}^{-1} \equiv (\alpha_1^{-1}, \alpha_2^{-1}, \ldots, \alpha_n^{-1}).$$

If $\alpha_i > 0$ and $\alpha_i^{-1} > 0$ for $i = 1, 2, \ldots, n$ then

$$P(\tilde{\alpha}\cdot\mathbf{p^0}, \tilde{\alpha}\cdot\mathbf{p^1}, \tilde{\alpha}^{-1}\cdot\mathbf{x^0}, \tilde{\alpha}^{-1}\cdot\mathbf{x^1}) = P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}).$$

1. Ideal Index

$$\begin{aligned} P^{id}(\tilde{\alpha}\cdot\mathbf{p^0}, \tilde{\alpha}\cdot\mathbf{p^1}, \tilde{\alpha}^{-1}\cdot\mathbf{x^0}, \tilde{\alpha}^{-1}\cdot\mathbf{x^1}) &= \sqrt{\frac{(\tilde{\alpha}\mathbf{p_I^1}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^0})(\tilde{\alpha}\mathbf{p_I^1}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^1})}{(\tilde{\alpha}\mathbf{p_I^0}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^0})(\tilde{\alpha}\mathbf{p_I^0}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^1})}} \\ &= \sqrt{\frac{\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^1}\cdot\mathbf{x_I^0})\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^0}\cdot\mathbf{x_I^0})\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}} \\ &= \sqrt{\frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}} = P^{id}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}). \end{aligned}$$

2. Paasche Index

$$\begin{aligned} P^{pa}(\tilde{\alpha}\cdot\mathbf{p^0}, \tilde{\alpha}\cdot\mathbf{p^1}, \tilde{\alpha}^{-1}\cdot\mathbf{x^0}, \tilde{\alpha}^{-1}\cdot\mathbf{x^1}) &= \frac{(\tilde{\alpha}\mathbf{p_I^1}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^1})}{(\tilde{\alpha}\mathbf{p_I^0}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^1})} = \frac{\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} \\ &= \frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} = P^{pa}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}). \end{aligned}$$

3. Laspeyres Index

$$\begin{aligned} P^{la}(\tilde{\alpha}\cdot\mathbf{p^0}, \tilde{\alpha}\cdot\mathbf{p^1}, \tilde{\alpha}^{-1}\cdot\mathbf{x^0}, \tilde{\alpha}^{-1}\cdot\mathbf{x^1}) &= \frac{(\tilde{\alpha}\mathbf{p_I^0}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^0})}{(\tilde{\alpha}\mathbf{p_I^1}\cdot\tilde{\alpha}^{-1}\mathbf{x_I^0})} = \frac{\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^0}\cdot\mathbf{x_I^0})}{\tilde{\alpha}\tilde{\alpha}^{-1}(\mathbf{p_I^1}\cdot\mathbf{x_I^0})} \\ &= \frac{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})}{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})} = P^{la}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1}). \end{aligned}$$

4. Törnqvist – Theil Index

$$P^{tt}\left(\tilde{\alpha}\cdot\mathbf{p^0},\tilde{\alpha}\cdot\mathbf{p^1},\tilde{\alpha}^{-1}\cdot\mathbf{x^0},\tilde{\alpha}^{-1}\cdot\mathbf{x^1}\right)=\prod_{i\in I}\left(\frac{\tilde{\alpha}p_i^1}{\tilde{\alpha}p_i^0}\right)^{.5\left(\frac{\tilde{\alpha}p_i^1\tilde{\alpha}^{-1}x_i^1}{\tilde{\alpha}\mathbf{p^1}\cdot\tilde{\alpha}^{-1}\mathbf{x^1}}+\frac{\tilde{\alpha}p_i^0\tilde{\alpha}^{-1}x_i^0}{\tilde{\alpha}\mathbf{p^0}\cdot\tilde{\alpha}^{-1}\mathbf{x^0}}\right)}$$

$$=\prod_{i\in I}\left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{\tilde{\alpha}\tilde{\alpha}^{-1}p_i^1x_i^1}{\tilde{\alpha}\tilde{\alpha}^{-1}\left(\mathbf{p^1}\cdot\mathbf{x^1}\right)}+\frac{\tilde{\alpha}\tilde{\alpha}^{-1}p_i^0x_i^0}{\tilde{\alpha}\tilde{\alpha}^{-1}\left(\mathbf{p^0}\cdot\mathbf{x^0}\right)}\right)}$$

$$=\prod_{i\in I}\left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{p_i^1x_i^1}{\mathbf{p^1}\cdot\mathbf{x^1}}+\frac{p_i^0x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}\right)}=P^{tt}\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right).$$

<u>*Axiom 5 (Consumer Reversal):*</u>

$$P\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)=\frac{1}{P\left(\mathbf{p^1},\mathbf{p^0},\mathbf{x^1},\mathbf{x^0}\right)}.$$

1. Ideal Index

$$P^{id}\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)=\sqrt{\frac{\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}}=\frac{1}{\sqrt{\frac{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)}{\left(\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)}}}=\frac{1}{P^{id}\left(\mathbf{p^1},\mathbf{p^0},\mathbf{x^1},\mathbf{x^0}\right)}.$$

2. Paasche Index – Consumer Reversal Not Satisfied

$$P^{pa}\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)=\frac{\left(\mathbf{p_I^1}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}\neq\frac{1}{\frac{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)}{\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)}}=\frac{1}{P^{pa}\left(\mathbf{p^1},\mathbf{p^0},\mathbf{x^1},\mathbf{x^0}\right)}.$$

3. Laspeyres Index – Consumer Reversal Not Satisfied

$$P^{la}\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)=\frac{\left(\mathbf{p_I^1}\cdot\mathbf{x_I^0}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^0}\right)}\neq\frac{1}{\frac{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}{\left(\mathbf{p_I^0}\cdot\mathbf{x_I^1}\right)}}=\frac{1}{P^{la}\left(\mathbf{p^1},\mathbf{p^0},\mathbf{x^1},\mathbf{x^0}\right)}.$$

4. Törnqvist – Theil Index

$$P^{tt}\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)=\prod_{i\in I}\left(\frac{p_i^0}{p_i^1}\right)^{.5\left(\frac{p_i^0 x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_i^1 x_i^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$$

$$=\frac{1}{\prod_{i\in I}\left(\frac{p_i^1}{p_i^0}\right)^{.5\left(\frac{p_i^1 x_i^1}{\mathbf{p^1}\cdot\mathbf{x^1}}+\frac{p_i^0 x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}\right)}}=\frac{1}{P^{tt}\left(\mathbf{p^1},\mathbf{p^0},\mathbf{x^1},\mathbf{x^0}\right)}.$$

<u>*Axiom 6 (Commodity Reversal):*</u> Define $\mathbf{p}^{\hat{\mathbf{i}}}$ as a permutation of $\mathbf{p^i}$ and $\mathbf{x}^{\hat{\mathbf{i}}}$ as the same permutation of $\mathbf{x^i}$ for **i = 0, 1**. Then

$$P\left(\mathbf{p}^{\hat{\mathbf{0}}},\mathbf{p}^{\hat{\mathbf{1}}},\mathbf{x}^{\hat{\mathbf{0}}},\mathbf{x}^{\hat{\mathbf{1}}}\right)=P\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)$$

This axiom clearly holds for the Ideal, Paasche, and Laspeyres indices given the distributive property of addition [see equations (3.6) – (3.8)]. Since multiplication is also distributive, this axiom also holds for the Törnqvist – Theil index [see equation (3.9)].

<u>*Axiom 7 (Monotonicity):*</u> If at least one of the commodity prices faced by the observed consumer decreases (increases) then the corresponding group price must also decrease (increase), ceteris paribus. Mathematically,

$$P\left(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}\right)\geq P\left(\mathbf{p^0},\mathbf{p}^{\hat{\mathbf{1}}},\mathbf{x^0},\mathbf{x^1}\right)\text{where } p_i^{\hat{1}}\leq p_i^1\ \forall\, i\in I.$$

Define $\mathbf{p^0}, \mathbf{p^1}$, and $\mathbf{p}^{\hat{\mathbf{1}}}$ by

$$\mathbf{p^0}\equiv\left(p_1^0, p_2^0,\ldots,p_n^0\right)$$
$$\mathbf{p^1}\equiv\left(p_1^1, p_2^1,\ldots,p_n^1\right)$$
$$\mathbf{p}^{\hat{\mathbf{1}}}\equiv\left(p_1^{\hat{1}}, p_2^{\hat{1}},\ldots,p_n^{\hat{1}}\right)$$

where and $p_i^{\hat{1}}\leq p_i^1\ \forall\, i\neq 1$. Define $\mathbf{x^0}, \mathbf{x^1}$, and $\mathbf{x}^{\hat{\mathbf{1}}}$ as corresponding consumption vectors and assume $\mathbf{x^1}=\mathbf{x}^{\hat{\mathbf{1}}}$.

1. Ideal Index

It is shown in 2 and 3 below that $\frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} \geq \frac{(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}$ and $\frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})} \geq \frac{(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^0})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})}$.

$$\Rightarrow \frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} \geq \frac{(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^0})(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} \Rightarrow \sqrt{\frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}} \geq \sqrt{\frac{(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^0})(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})(\mathbf{p_I^0}\cdot\mathbf{x_I^1})}}$$

$$\Rightarrow P^{id}(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}) \geq P^{id}(\mathbf{p^0},\mathbf{p^{\hat{1}}},\mathbf{x^0},\mathbf{x^1})$$

2. Paasche Index

By assumption $p_i^{\hat{1}} \geq p_i^1 \; \forall \, i$ and $\mathbf{x^1} = \mathbf{x^{\hat{1}}}$.

$$\Rightarrow p_1^1 x_1^1 + \cdots + p_n^1 x_n^1 \geq p_1^{\hat{1}} x_1^1 + \cdots + p_n^{\hat{1}} x_n^1 \Rightarrow (\mathbf{p_I^1}\cdot\mathbf{x_I^1}) \geq (\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^1})$$

$$\Rightarrow \frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} \geq \frac{(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^1})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^1})} \Rightarrow P^{pa}(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}) \geq P^{pa}(\mathbf{p^0},\mathbf{p^{\hat{1}}},\mathbf{x^0},\mathbf{x^1})$$

3. Laspeyres Index

By assumption $p_i^1 \geq p_i^{\hat{1}} \forall \, i$ and $\mathbf{x^1} = \mathbf{x^{\hat{1}}}$.

$$\Rightarrow p_1^1 x_1^0 + \cdots + p_n^1 x_n^0 \geq p_1^{\hat{1}} x_1^0 + \cdots + p_n^{\hat{1}} x_n^0 \Rightarrow (\mathbf{p_I^1}\cdot\mathbf{x}_\mathbf{I}^0) \geq (\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x}_\mathbf{I}^0)$$

$$\Rightarrow \frac{(\mathbf{p_I^1}\cdot\mathbf{x_I^0})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})} \geq \frac{(\mathbf{p_I^{\hat{1}}}\cdot\mathbf{x_I^0})}{(\mathbf{p_I^0}\cdot\mathbf{x_I^0})} \Rightarrow P^{la}(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}) \geq P^{la}(\mathbf{p^0},\mathbf{p^{\hat{1}}},\mathbf{x^0},\mathbf{x^1})$$

4. Törnqvist – Theil Index – Monotonicity Not Satisfied

It is easiest to show the TT index fails the monotonicity axiom using a counterexample.

Consider a two good case where $\mathbf{p^0} \equiv (1000,1)$, $\mathbf{p^1} \equiv (1000,1000)$, $\mathbf{p^{\hat{1}}} \equiv (200,1000)$ and $\mathbf{x^0} = \mathbf{x^1} \equiv (1,1)$. In this example, $p_1^{\hat{1}} = 200 < p_1^1 = 1000$ and $p_2^{\hat{1}} = p_2^1$.
However, $P^{tt}(\mathbf{p^0},\mathbf{p^1},\mathbf{x^0},\mathbf{x^1}) = 5.64 < P^{tt}(\mathbf{p^0},\mathbf{p^{\hat{1}}},\mathbf{x^0},\mathbf{x^1}) = 6.98$, which is a violation of monotonicity.

Axiom 8 (Mean Value Test): Group prices must lie between the smallest and largest ratios of commodity prices between the observed and base consumers. Mathematically,

$$\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P \le \max_i\left(\frac{p_i^1}{p_i^0}\right).$$

For each index, assume $\min_i\left(\frac{p_i^1}{p_i^0}\right) = \frac{p_1^1}{p_1^0}$ and $\max_i\left(\frac{p_i^1}{p_i^0}\right) = \frac{p_2^1}{p_2^0}$. This implies $\frac{p_1^1}{p_1^0} \le \frac{p_i^1}{p_i^0}$ and $\frac{p_2^1}{p_2^0} \ge \frac{p_i^1}{p_i^0}\ \forall i.$

Each of the following proofs is broken into two parts:

a) show that $\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P$, and

b) show that $\max_i\left(\frac{p_i^1}{p_i^0}\right) \ge P$.

1. Ideal Index

a) It is shown in (2) and (3) below that $\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P^{pa}$ and $\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P^{la}$. Thus,

$$\left(\min_i\left(\frac{p_i^1}{p_i^0}\right)\right)^2 \le P^{pa}P^{la} \Rightarrow \min_i\left(\frac{p_i^1}{p_i^0}\right) \le \sqrt{P^{pa}P^{la}} = P^{id}.$$

b) It is shown in (2) and (3) below that $\max_i\left(\frac{p_i^1}{p_i^0}\right) \ge P^{pa}$ and $\max_i\left(\frac{p_i^1}{p_i^0}\right) \ge P^{la}$. Thus,

$$\left(\max_i\left(\frac{p_i^1}{p_i^0}\right)\right)^2 \ge P^{pa}P^{la} \Rightarrow \max_i\left(\frac{p_i^1}{p_i^0}\right) \ge \sqrt{P^{pa}P^{la}} = P^{id}.$$

2. Paasche Index

a) Assume $\min_i\left(\frac{p_i^1}{p_i^0}\right) > P^{pa}$. $\Rightarrow \frac{p_1^1}{p_1^0} > \frac{\mathbf{p_I^1} \cdot \mathbf{x_I^1}}{\mathbf{p_I^0} \cdot \mathbf{x_I^1}} = \frac{p_1^1 x_1^1 + p_2^1 x_2^1 + \cdots + p_n^1 x_n^1}{p_1^0 x_1^1 + p_2^0 x_2^1 + \cdots + p_n^0 x_n^1}$

$\Rightarrow \frac{p_1^0 x_1^1}{p_1^0} + \frac{p_2^0 x_2^1}{p_1^0} + \cdots + \frac{p_n^0 x_n^1}{p_1^0} > \frac{p_1^1 x_1^1}{p_1^1} + \frac{p_2^1 x_2^1}{p_1^1} + \cdots + \frac{p_n^1 x_n^1}{p_1^1}$

By assumption, $\frac{p_1^1}{p_1^0} \le \frac{p_i^1}{p_i^0} \forall i \Rightarrow p_1^1 \le \frac{p_1^0 p_i^1}{p_i^0} \forall i.$ Substituting into the above equation and simplifying yields $\frac{p_2^0 x_2^1}{p_1^0} + \frac{p_3^0 x_3^1}{p_1^0} + \cdots + \frac{p_n^0 x_n^1}{p_1^0} > \frac{p_2^0 x_2^1}{p_1^0} + \frac{p_3^0 x_3^1}{p_1^0} + \cdots + \frac{p_n^0 x_n^1}{p_1^0}$ which is clearly not true. Thus, by contradiction, $\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P^{pa}$.

b) Assume $\max_i\left(\frac{p_i^1}{p_i^0}\right) < P^{pa}$. $\Rightarrow \frac{p_2^1}{p_2^0} < \frac{\mathbf{p_I^1} \cdot \mathbf{x_I^1}}{\mathbf{p_I^0} \cdot \mathbf{x_I^1}} = \frac{p_1^1 x_1^1 + p_2^1 x_2^1 + \cdots + p_n^1 x_n^1}{p_1^0 x_1^1 + p_2^0 x_2^1 + \cdots + p_n^0 x_n^1}$

$\Rightarrow \frac{p_1^0 x_1^1}{p_2^0} + \frac{p_2^0 x_2^1}{p_2^0} + \cdots + \frac{p_n^0 x_n^1}{p_2^0} < \frac{p_1^1 x_1^1}{p_2^1} + \frac{p_2^1 x_2^1}{p_2^1} + \cdots + \frac{p_n^1 x_n^1}{p_2^1}$

By assumption, $\frac{p_2^1}{p_2^0} \ge \frac{p_i^1}{p_i^0} \forall i \Rightarrow p_2^1 \ge \frac{p_2^0 p_i^1}{p_i^0} \forall i.$ Substituting into the above equation and simplifying yields $\frac{p_1^0 x_1^1}{p_2^0} + \frac{p_3^0 x_3^1}{p_2^0} + \frac{p_4^0 x_4^1}{p_2^0} + \cdots + \frac{p_n^0 x_n^1}{p_2^0} > \frac{p_1^0 x_1^1}{p_2^0} + \frac{p_3^0 x_3^1}{p_2^0} + \frac{p_4^0 x_4^1}{p_2^0} + \cdots + \frac{p_n^0 x_n^1}{p_2^0}$ which is clearly not true. Thus, by contradiction, $\max_i\left(\frac{p_i^1}{p_i^0}\right) \ge P^{pa}$.

3. Laspeyres Index

a) Assume $\min_i\left(\frac{p_i^1}{p_i^0}\right) > P^{la}$. $\Rightarrow \frac{p_1^1}{p_1^0} > \frac{\mathbf{p_I^1} \cdot \mathbf{x_I^0}}{\mathbf{p_I^0} \cdot \mathbf{x_I^0}} = \frac{p_1^1 x_1^0 + p_2^1 x_2^0 + \cdots + p_n^1 x_n^0}{p_1^0 x_1^0 + p_2^0 x_2^0 + \cdots + p_n^0 x_n^0}$

$\Rightarrow \frac{p_1^0 x_1^0}{p_1^0} + \frac{p_2^0 x_2^0}{p_1^0} + \cdots + \frac{p_n^0 x_n^0}{p_1^0} > \frac{p_1^1 x_1^0}{p_1^1} + \frac{p_2^1 x_2^0}{p_1^1} + \cdots + \frac{p_n^1 x_n^0}{p_1^1}$

By assumption, $\frac{p_1^1}{p_1^0} \le \frac{p_i^1}{p_i^0} \forall i \Rightarrow p_1^1 \le \frac{p_1^0 p_i^1}{p_i^0} \forall i.$ Substituting into the above equation and simplifying yields $\frac{p_2^0 x_2^0}{p_1^0} + \frac{p_3^0 x_3^0}{p_1^0} + \cdots + \frac{p_n^0 x_n^0}{p_1^0} > \frac{p_2^0 x_2^0}{p_1^0} + \frac{p_3^0 x_3^0}{p_1^0} + \cdots + \frac{p_n^0 x_n^0}{p_1^0}$

which is clearly not true. Thus, by contradiction, $\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P^{la}$.

b) Assume $\max_i\left(\frac{p_i^1}{p_i^0}\right) < P^{la}$. $\Rightarrow \frac{p_2^1}{p_2^0} < \frac{\mathbf{p_I^1} \cdot \mathbf{x_I^0}}{\mathbf{p_I^0} \cdot \mathbf{x_I^0}} = \frac{p_1^1 x_1^0 + p_2^1 x_2^0 + \cdots + p_n^1 x_n^0}{p_1^0 x_1^0 + p_2^0 x_2^0 + \cdots + p_n^0 x_n^0}$

$\Rightarrow \frac{p_1^0 x_1^0}{p_2^0} + \frac{p_2^0 x_2^0}{p_2^0} + \cdots + \frac{p_n^0 x_n^0}{p_2^0} < \frac{p_1^1 x_1^0}{p_2^1} + \frac{p_2^1 x_2^0}{p_2^1} + \cdots + \frac{p_n^1 x_n^0}{p_2^1}$

By assumption, $\frac{p_2^1}{p_2^0} \ge \frac{p_i^1}{p_i^0} \forall i \Rightarrow p_2^1 \ge \frac{p_2^0 p_i^1}{p_i^0} \forall i.$ Substituting into the above equation and simplifying yields $\frac{p_1^0 x_1^1}{p_2^0} + \frac{p_3^0 x_3^1}{p_2^0} + \frac{p_4^0 x_4^1}{p_2^0} + \cdots + \frac{p_n^0 x_n^1}{p_2^0} < \frac{p_1^0 x_1^1}{p_2^0} + \frac{p_3^0 x_3^1}{p_2^0} + \frac{p_4^0 x_4^1}{p_2^0} + \cdots + \frac{p_n^0 x_n^1}{p_2^0}$

which is clearly not true. Thus, by contradiction, $\max_i\left(\frac{p_i^1}{p_i^0}\right) \ge P^{pa}$.

4. Törnqvist – Theil

a) Assume $\min_i\left(\frac{p_i^1}{p_i^0}\right) > P^{tt}$. $\Rightarrow \frac{p_1^1}{p_1^0} > \prod_{i\in I}\left(\frac{p_i^0}{p_i^1}\right)^{.5\left(\frac{p_i^0 x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_i^1 x_i^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$

$$\Rightarrow \frac{p_1^1}{p_1^0} > \left(\frac{p_1^0}{p_1^1}\right)^{.5\left(\frac{p_1^0 x_1^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_1^1 x_1^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)} * \left(\frac{p_2^0}{p_2^1}\right)^{.5\left(\frac{p_2^0 x_2^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_2^1 x_2^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)} * \ldots * \left(\frac{p_n^0}{p_n^1}\right)^{.5\left(\frac{p_n^0 x_n^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_n^1 x_n^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$$

By assumption, $\frac{p_1^1}{p_1^0} \le \frac{p_i^1}{p_i^0}\ \forall i$. Substituting into the above equation and simplifying yields

$$\frac{p_1^1}{p_1^0} > \left(\frac{p_1^0}{p_1^1}\right)^{.5\left(\frac{p_1^0 x_1^0 + p_2^0 x_2^0 + p_n^0 x_n^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_1^1 x_1^1 + p_2^1 x_2^1 + p_n^1 x_n^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$$

$$\Rightarrow \ln\left(\frac{p_1^0}{p_1^1}\right) > .5\left(\frac{p_1^0 x_1^0 + p_2^0 x_2^0 + p_n^0 x_n^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_1^1 x_1^1 + p_2^1 x_2^1 + p_n^1 x_n^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)\ln\left(\frac{p_1^0}{p_1^1}\right)$$

$\Rightarrow 2 > (1+1)$, which is clearly not true. Thus, by contradiction, $\min_i\left(\frac{p_i^1}{p_i^0}\right) \le P^{tt}$.

b) Assume $\max_i\left(\frac{p_i^1}{p_i^0}\right) < P^{tt}$. $\Rightarrow \frac{p_2^1}{p_2^0} < \prod_{i\in I}\left(\frac{p_i^0}{p_i^1}\right)^{.5\left(\frac{p_i^0 x_i^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_i^1 x_i^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$

$$\Rightarrow \frac{p_2^1}{p_2^0} < \left(\frac{p_1^0}{p_1^1}\right)^{.5\left(\frac{p_1^0 x_1^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_1^1 x_1^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)} * \left(\frac{p_2^0}{p_2^1}\right)^{.5\left(\frac{p_2^0 x_2^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_2^1 x_2^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)} * \ldots * \left(\frac{p_n^0}{p_n^1}\right)^{.5\left(\frac{p_n^0 x_n^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_n^1 x_n^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$$

By assumption, $\frac{p_2^1}{p_2^0} \ge \frac{p_i^1}{p_i^0}\ \forall i$. Substituting into the above equation and simplifying yields

$$\frac{p_2^1}{p_2^0} < \left(\frac{p_2^0}{p_2^1}\right)^{.5\left(\frac{p_1^0 x_1^0 + p_2^0 x_2^0 + p_n^0 x_n^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_1^1 x_1^1 + p_2^1 x_2^1 + p_n^1 x_n^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)}$$

$$\Rightarrow \ln\left(\frac{p_2^0}{p_2^1}\right) < .5\left(\frac{p_1^0 x_1^0 + p_2^0 x_2^0 + p_n^0 x_n^0}{\mathbf{p^0}\cdot\mathbf{x^0}}+\frac{p_1^1 x_1^1 + p_2^1 x_2^1 + p_n^1 x_n^1}{\mathbf{p^1}\cdot\mathbf{x^1}}\right)\ln\left(\frac{p_2^0}{p_2^1}\right)$$

$\Rightarrow 2 < (1+1)$, which is clearly not true. Thus, by contradiction, $\max_i\left(\frac{p_i^1}{p_i^0}\right) \ge P^{tt}$.

Axiom 9 (Circularity): Circularity is a transitivity condition that is more relevant to time series data. Define $\mathbf{p^2}$ as the price vector observed in period 2. Define $\mathbf{p^1}$ as the price vector observed in period 1. Define $\mathbf{p^0}$ as the price vector observed in the base period. Define $\mathbf{x^2}$, $\mathbf{x^1}$ and $\mathbf{x^0}$ as the consumption vectors corresponding to $\mathbf{p^2}$, $\mathbf{p^1}$ and $\mathbf{p^0}$ respectively. Circularity requires

$$P(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1})P(\mathbf{p^1}, \mathbf{p^2}, \mathbf{x^1}, \mathbf{x^2}) = P(\mathbf{p^0}, \mathbf{p^2}, \mathbf{x^0}, \mathbf{x^2}).$$

It is easiest to show that each of the four indices considered here fail the circularity axiom by counterexample. For each of the following counterexamples, assume a two good case where $\mathbf{p^0}$ *= (2, 3)*, $\mathbf{p^1}$ *= (1, 2)*, $\mathbf{p^2}$ *= (3, 4)*, $\mathbf{x^0}$ **=** *(2, 5)*, and $\mathbf{x^1}$ *= (1, 3) =* $\mathbf{x^2}$ *= (3, 4)*.

1. Ideal Index – Circularity Not Satisfed

$$P^{id}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1})P^{id}(\mathbf{p^1}, \mathbf{p^2}, \mathbf{x^1}, \mathbf{x^2}) = 1.45 \neq P^{id}(\mathbf{p^0}, \mathbf{p^2}, \mathbf{x^0}, \mathbf{x^2}) = 1.39.$$

2. Paasche Index – Circularity Not Satisfed

$$P^{pa}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1})P^{id}(\mathbf{p^1}, \mathbf{p^2}, \mathbf{x^1}, \mathbf{x^2}) = 1.55 \neq P^{pa}(\mathbf{p^0}, \mathbf{p^2}, \mathbf{x^0}, \mathbf{x^2}) = 1.42.$$

3. Laspeyres Index – Circularity Not Satisfed

$$P^{la}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1})P^{la}(\mathbf{p^1}, \mathbf{p^2}, \mathbf{x^1}, \mathbf{x^2}) = 1.35 \neq P^{la}(\mathbf{p^0}, \mathbf{p^2}, \mathbf{x^0}, \mathbf{x^2}) = 1.37.$$

4. Törnqvist – Theil Index – Circularity Not Satisfed

$$P^{tt}(\mathbf{p^0}, \mathbf{p^1}, \mathbf{x^0}, \mathbf{x^1})P^{tt}(\mathbf{p^1}, \mathbf{p^2}, \mathbf{x^1}, \mathbf{x^2}) = 1.45 \neq P^{tt}(\mathbf{p^0}, \mathbf{p^2}, \mathbf{x^0}, \mathbf{x^2}) = 1.39.$$

APPENDIX B: SAS CODE

The following SAS code was used to derive the parameter and elasticity estimates reported in Tables 4.5 – 4.7. In addition, it was used to calculate the correlation coefficients used to determine the number of violations of the Generalized Composite Commodity Theorem for each of the regions under data partition I (see Tables 4.2 – 4.4). The code used to determine the number of violations of the GCC for partitions II and III was slightly different. However, the differences are minor and only involve changing the way in which the groups are created. Therefore, these variations are omitted.

```
/* this block of code reads in the price and expenditure data. */

data ce2;
infile 'd:\thesis\ce.txt' DLM='09'x;

input FAMNUM $ BNO1ST1  BNO1WT1  BNO1PR1  BNO1ST2  BNO1WT2  BNO1PR2
BNO1ST3  BNO1WT3  BNO1PR3  BNO1WT4  BNO1PR4  BNO1WT5  BNO1PR5
BNO1WT6  BNO1PR6  BNO1WT7  BNO1PR7  BNO1WT8  BNO1PR8  BNO1WT9
BNO1PR9  BNO1WT10 BNO1PR10 BNO2ST1  BNO2WT1  BNO2PR1  BNO2ST2
BNO2WT2  BNO2PR2  BNO2ST3  BNO2WT3  BNO2PR3  BNO3ST1  BNO3WT1
BNO3PR1  BNO3ST2  BNO3WT2  BNO3PR2  BNO3ST3  BNO3WT3  BNO3PR3
BNO3WT4  BNO3PR4  BNO4ST1  BNO4WT1  BNO4PR1  BNO5ST1  BNO5WT1
BNO5PR1  BNO5ST2  BNO5WT2  BNO5PR2  BNO5ST3  BNO5WT3  BNO5PR3
BNO5WT4  BNO5PR4  BNO5WT5  BNO5PR5  BNO6ST1  BNO6WT1  BNO6PR1
BNO6ST2  BNO6WT2  BNO6PR2  BNO6ST3  BNO6WT3  BNO6PR3  BNO6WT4
BNO6PR4  BNO7ST1  BNO7WT1  BNO7PR1  BNO7ST2  BNO7WT2  BNO7PR2
BNO8ST1  BNO8WT1  BNO8PR1  BNO8ST2  BNO8WT2  BNO8PR2  BNO9ST1
BNO9WT1  BNO9PR1  BNO9ST2  BNO9WT2  BNO9PR2  BNO9ST3  BNO9WT3
BNO9PR3  BNO9WT4  BNO9PR4  BNO9WT5  BNO9PR5  BNO9WT6  BNO9PR6
BNO9WT7  BNO9PR7  BDE1ST1  BDE1WT1  BDE1PR1  BDE1ST2  BDE1WT2
BDE1PR2  BDE1ST3  BDE1WT3  BDE1PR3  BDE1WT4  BDE1PR4  BDE1WT5
BDE1PR5  BDE1WT6  BDE1PR6  BDE1WT7  BDE1PR7  BDE1WT8  BDE1PR8
BDE1WT9  BDE1PR9  BDE2ST1  BDE2WT1  BDE2PR1  BDE2ST2  BDE2WT2
BDE2PR2  BDE2ST3  BDE2WT3  BDE2PR3  BDE3ST1  BDE3WT1  BDE3PR1
BDE3ST2  BDE3WT2  BDE3PR2  BDE3ST3  BDE3WT3  BDE3PR3  BDE4ST1
BDE4WT1  BDE4PR1  BDE4ST2  BDE4WT2  BDE4PR2  BDE5ST1  BDE5WT1
BDE5PR1  BDE5ST2  BDE5WT2  BDE5PR2  BDE5ST3  BDE5WT3  BDE5PR3
BDE5WT4  BDE5PR4  BDE5WT5  BDE5PR5  BDE5WT6  BDE5PR6  BDE5WT7
BDE5PR7  BDE5WT8  BDE5PR8  BDE6ST1  BDE6WT1  BDE6PR1  BDE6ST2
BDE6WT2  BDE6PR2  BDE6ST3  BDE6WT3  BDE6PR3  BDE6WT4  BDE6PR4
BDE7ST1  BDE7WT1  BDE7PR1  BDE7ST2  BDE7WT2  BDE7PR2  BDE8ST1
BDE8WT1  BDE8PR1  BDE8ST2  BDE8WT2  BDE8PR2  BDE9ST1  BDE9WT1
BDE9PR1  BDE9ST2  BDE9WT2  BDE9PR2  BDE9ST3  BDE9WT3  BDE9PR3
BDE9WT4  BDE9PR4  BJA1ST1  BJA1WT1  BJA1PR1  BJA1ST2  BJA1WT2
BJA1PR2  BJA1ST3  BJA1WT3  BJA1PR3  BJA1WT4  BJA1PR4  BJA1WT5
BJA1PR5  BJA1WT6  BJA1PR6  BJA1WT7  BJA1PR7  BJA1WT8  BJA1PR8
BJA1WT9  BJA1PR9  BJA1WT10 BJA1PR10 BJA2ST1  BJA2WT1  BJA2PR1
BJA2ST2  BJA2WT2  BJA2PR2  BJA2ST3  BJA2WT3  BJA2PR3  BJA2WT4
BJA2PR4  BJA3ST1  BJA3WT1  BJA3PR1  BJA3ST2  BJA3WT2  BJA3PR2
BJA3ST3  BJA3WT3  BJA3PR3  BJA4ST1  BJA4WT1  BJA4PR1  BJA5ST1
```

BJA5WT1 BJA5PR1 BJA5ST2 BJA5WT2 BJA5PR2 BJA5ST3 BJA5WT3
BJA5PR3 BJA5WT4 BJA5PR4 BJA5WT5 BJA5PR5 BJA6ST1 BJA6WT1
BJA6PR1 BJA6ST2 BJA6WT2 BJA6PR2 BJA6ST3 BJA6WT3 BJA6PR3
BJA6WT4 BJA6PR4 BJA7ST1 BJA7WT1 BJA7PR1 BJA7ST2 BJA7WT2
BJA7PR2 BJA8ST1 BJA8WT1 BJA8PR1 BJA8ST2 BJA8WT2 BJA8PR2
BJA8ST3 BJA8WT3 BJA8PR3 BJA9ST1 BJA9WT1 BJA9PR1 BJA9ST2
BJA9WT2 BJA9PR2 BJA9ST3 BJA9WT3 BJA9PR3 BJA9WT4 BJA9PR4
BFE1ST1 BFE1WT1 BFE1PR1 BFE1ST2 BFE1WT2 BFE1PR2 BFE1ST3
BFE1WT3 BFE1PR3 BFE1WT4 BFE1PR4 BFE1WT5 BFE1PR5 BFE1WT6
BFE1PR6 BFE1WT7 BFE1PR7 BFE1WT8 BFE1PR8 BFE2ST1 BFE2WT1
BFE2PR1 BFE2ST2 BFE2WT2 BFE2PR2 BFE2ST3 BFE2WT3 BFE2PR3
BFE3ST1 BFE3WT1 BFE3PR1 BFE3ST2 BFE3WT2 BFE3PR2 BFE3ST3
BFE3WT3 BFE3PR3 BFE3WT4 BFE3PR4 BFE3WT5 BFE3PR5 BFE4ST1
BFE4WT1 BFE4PR1 BFE5ST1 BFE5WT1 BFE5PR1 BFE5ST2 BFE5WT2
BFE5PR2 BFE5ST3 BFE5WT3 BFE5PR3 BFE5WT4 BFE5PR4 BFE5WT5
BFE5PR5 BFE5WT6 BFE5PR6 BFE5WT7 BFE5PR7 BFE5WT8 BFE5PR8
BFE5WT9 BFE5PR9 BFE6ST1 BFE6WT1 BFE6PR1 BFE6ST2 BFE6WT2
BFE6PR2 BFE6ST3 BFE6WT3 BFE6PR3 BFE6WT4 BFE6PR4 BFE7ST1
BFE7WT1 BFE7PR1 BFE7ST2 BFE7WT2 BFE7PR2 BFE8ST1 BFE8WT1
BFE8PR1 BFE8ST2 BFE8WT2 BFE8PR2 BFE8ST3 BFE8WT3 BFE8PR3
BFE9ST1 BFE9WT1 BFE9PR1 BFE9ST2 BFE9WT2 BFE9PR2 BFE9ST3
BFE9WT3 BFE9PR3 BFE9WT4 BFE9PR4 BMR1ST1 BMR1WT1 BMR1PR1
BMR1ST2 BMR1WT2 BMR1PR2 BMR1ST3 BMR1WT3 BMR1PR3 BMR1WT4
BMR1PR4 BMR1WT5 BMR1PR5 BMR1WT6 BMR1PR6 BMR1WT7 BMR1PR7
BMR1WT8 BMR1PR8 BMR1WT9 BMR1PR9 BMR1WT10 BMR1PR10 BMR1WT11
BMR1PR11 BMR1WT12 BMR1PR12 BMR2ST1 BMR2WT1 BMR2PR1 BMR2ST2
BMR2WT2 BMR2PR2 BMR2ST3 BMR2WT3 BMR2PR3 BMR3ST1 BMR3WT1
BMR3PR1 BMR3ST2 BMR3WT2 BMR3PR2 BMR3ST3 BMR3WT3 BMR3PR3
BMR3WT4 BMR3PR4 BMR3WT5 BMR3PR5 BMR4ST1 BMR4WT1 BMR4PR1
BMR5ST1 BMR5WT1 BMR5PR1 BMR5ST2 BMR5WT2 BMR5PR2 BMR5ST3
BMR5WT3 BMR5PR3 BMR5WT4 BMR5PR4 BMR5WT5 BMR5PR5 BMR5WT6
BMR5PR6 BMR5WT7 BMR5PR7 BMR5WT8 BMR5PR8 BMR5WT9 BMR5PR9
BMR6ST1 BMR6WT1 BMR6PR1 BMR6ST2 BMR6WT2 BMR6PR2 BMR6ST3
BMR6WT3 BMR6PR3 BMR6WT4 BMR6PR4 BMR7ST1 BMR7WT1 BMR7PR1
BMR7ST2 BMR7WT2 BMR7PR2 BMR7ST3 BMR7WT3 BMR7PR3 BMR8ST1
BMR8WT1 BMR8PR1 BMR8ST2 BMR8WT2 BMR8PR2 BMR8ST3 BMR8WT3
BMR8PR3 BMR9ST1 BMR9WT1 BMR9PR1 BMR9ST2 BMR9WT2 BMR9PR2
BMR9ST3 BMR9WT3 BMR9PR3 BMR9WT4 BMR9PR4 CNO1ST1 CNO1WT1
CNO1PR1 CNO1ST2 CNO1WT2 CNO1PR2 CNO1ST3 CNO1WT3 CNO1PR3
CNO1WT4 CNO1PR4 CNO2ST1 CNO2WT1 CNO2PR1 CNO2ST2 CNO2WT2
CNO2PR2 CNO2ST3 CNO2WT3 CNO2PR3 CNO2WT4 CNO2PR4 CNO2WT5
CNO2PR5 CNO3ST1 CNO3WT1 CNO3PR1 CNO4ST1 CNO4WT1 CNO4PR1
CNO4ST2 CNO4WT2 CNO4PR2 CNO4ST3 CNO4WT3 CNO4PR3 CNO4WT4
CNO4PR4 CNO4WT5 CNO4PR5 CNO5ST1 CNO5WT1 CNO5PR1 CNO5ST2
CNO5WT2 CNO5PR2 CNO5ST3 CNO5WT3 CNO5PR3 CNO5WT4 CNO5PR4
CNO6ST1 CNO6WT1 CNO6PR1 CNO6ST2 CNO6WT2 CNO6PR2 CNO6ST3
CNO6WT3 CNO6PR3 CNO7ST1 CNO7WT1 CNO7PR1 CNO7ST2 CNO7WT2
CNO7PR2 CNO7ST3 CNO7WT3 CNO7PR3 CNO7WT4 CNO7PR4 CNO8ST1
CNO8WT1 CNO8PR1 CNO8ST2 CNO8WT2 CNO8PR2 CNO8ST3 CNO8WT3
CNO8PR3 CNO8WT4 CNO8PR4 CNO9ST1 CNO9WT1 CNO9PR1 CNO9ST2
CNO9WT2 CNO9PR2 CNO0ST1 CNO0WT1 CNO0PR1 CNO0ST2 CNO0WT2
CNO0PR2 CNO0ST3 CNO0WT3 CNO0PR3 CNO0WT4 CNO0PR4 CNO0WT5
CNO0PR5 CDE1ST1 CDE1WT1 CDE1PR1 CDE1ST2 CDE1WT2 CDE1PR2
CDE1ST3 CDE1WT3 CDE1PR3 CDE1WT4 CDE1PR4 CDE2ST1 CDE2WT1
CDE2PR1 CDE2ST2 CDE2WT2 CDE2PR2 CDE2ST3 CDE2WT3 CDE2PR3
CDE2WT4 CDE2PR4 CDE2WT5 CDE2PR5 CDE3ST1 CDE3WT1 CDE3PR1

CDE3ST2 CDE3WT2 CDE3PR2 CDE4ST1 CDE4WT1 CDE4PR1 CDE4ST2
CDE4WT2 CDE4PR2 CDE4ST3 CDE4WT3 CDE4PR3 CDE4WT4 CDE4PR4
CDE4WT5 CDE4PR5 CDE4WT6 CDE4PR6 CDE4WT7 CDE4PR7 CDE4WT8
CDE4PR8 CDE5ST1 CDE5WT1 CDE5PR1 CDE5ST2 CDE5WT2 CDE5PR2
CDE5ST3 CDE5WT3 CDE5PR3 CDE6ST1 CDE6WT1 CDE6PR1 CDE6ST2
CDE6WT2 CDE6PR2 CDE7ST1 CDE7WT1 CDE7PR1 CDE7ST2 CDE7WT2
CDE7PR2 CDE7ST3 CDE7WT3 CDE7PR3 CDE7WT4 CDE7PR4 CDE8ST1
CDE8WT1 CDE8PR1 CDE8ST2 CDE8WT2 CDE8PR2 CDE8ST3 CDE8WT3
CDE8PR3 CDE9ST1 CDE9WT1 CDE9PR1 CDE9ST2 CDE9WT2 CDE9PR2
CDE0ST1 CDE0WT1 CDE0PR1 CDE0ST2 CDE0WT2 CDE0PR2 CDE0ST3
CDE0WT3 CDE0PR3 CDE0WT4 CDE0PR4 CDE0WT5 CDE0PR5 CJA1ST1
CJA1WT1 CJA1PR1 CJA1ST2 CJA1WT2 CJA1PR2 CJA1ST3 CJA1WT3
CJA1PR3 CJA1WT4 CJA1PR4 CJA2ST1 CJA2WT1 CJA2PR1 CJA2ST2
CJA2WT2 CJA2PR2 CJA2ST3 CJA2WT3 CJA2PR3 CJA2WT4 CJA2PR4
CJA2WT5 CJA2PR5 CJA2WT6 CJA2PR6 CJA2WT7 CJA2PR7 CJA3ST1
CJA3WT1 CJA3PR1 CJA3ST2 CJA3WT2 CJA3PR2 CJA3ST3 CJA3WT3
CJA3PR3 CJA4ST1 CJA4WT1 CJA4PR1 CJA4ST2 CJA4WT2 CJA4PR2
CJA4ST3 CJA4WT3 CJA4PR3 CJA4WT4 CJA4PR4 CJA4WT5 CJA4PR5
CJA4WT6 CJA4PR6 CJA5ST1 CJA5WT1 CJA5PR1 CJA5ST2 CJA5WT2
CJA5PR2 CJA6ST1 CJA6WT1 CJA6PR1 CJA6ST2 CJA6WT2 CJA6PR2
CJA6ST3 CJA6WT3 CJA6PR3 CJA7ST1 CJA7WT1 CJA7PR1 CJA7ST2
CJA7WT2 CJA7PR2 CJA7ST3 CJA7WT3 CJA7PR3 CJA7WT4 CJA7PR4
CJA7WT5 CJA7PR5 CJA7WT6 CJA7PR6 CJA8ST1 CJA8WT1 CJA8PR1
CJA8ST2 CJA8WT2 CJA8PR2 CJA9ST1 CJA9WT1 CJA9PR1 CJA0ST1
CJA0WT1 CJA0PR1 CJA0ST2 CJA0WT2 CJA0PR2 CJA0ST3 CJA0WT3
CJA0PR3 CJA0WT4 CJA0PR4 CJA0WT5 CJA0PR5 CFE1ST1 CFE1WT1
CFE1PR1 CFE1ST2 CFE1WT2 CFE1PR2 CFE1ST3 CFE1WT3 CFE1PR3
CFE1WT4 CFE1PR4 CFE2ST1 CFE2WT1 CFE2PR1 CFE2ST2 CFE2WT2
CFE2PR2 CFE2ST3 CFE2WT3 CFE2PR3 CFE2WT4 CFE2PR4 CFE2WT5
CFE2PR5 CFE3ST1 CFE3WT1 CFE3PR1 CFE3ST2 CFE3WT2 CFE3PR2
CFE3ST3 CFE3WT3 CFE3PR3 CFE3WT4 CFE3PR4 CFE4ST1 CFE4WT1
CFE4PR1 CFE4ST2 CFE4WT2 CFE4PR2 CFE4ST3 CFE4WT3 CFE4PR3
CFE4WT4 CFE4PR4 CFE4WT5 CFE4PR5 CFE4WT6 CFE4PR6 CFE5ST1
CFE5WT1 CFE5PR1 CFE5ST2 CFE5WT2 CFE5PR2 CFE6ST1 CFE6WT1
CFE6PR1 CFE6ST2 CFE6WT2 CFE6PR2 CFE7ST1 CFE7WT1 CFE7PR1
CFE7ST2 CFE7WT2 CFE7PR2 CFE7ST3 CFE7WT3 CFE7PR3 CFE7WT4
CFE7PR4 CFE8ST1 CFE8WT1 CFE8PR1 CFE8ST2 CFE8WT2 CFE8PR2
CFE8ST3 CFE8WT3 CFE8PR3 CFE9ST1 CFE9WT1 CFE9PR1 CFE0ST1
CFE0WT1 CFE0PR1 CFE0ST2 CFE0WT2 CFE0PR2 CFE0ST3 CFE0WT3
CFE0PR3 CFE0WT4 CFE0PR4 CMR1ST1 CMR1WT1 CMR1PR1 CMR1ST2
CMR1WT2 CMR1PR2 CMR1ST3 CMR1WT3 CMR1PR3 CMR1WT4 CMR1PR4
CMR1WT5 CMR1PR5 CMR2ST1 CMR2WT1 CMR2PR1 CMR2ST2 CMR2WT2
CMR2PR2 CMR2ST3 CMR2WT3 CMR2PR3 CMR2WT4 CMR2PR4 CMR2WT5
CMR2PR5 CMR2WT6 CMR2PR6 CMR3ST1 CMR3WT1 CMR3PR1 CMR3ST2
CMR3WT2 CMR3PR2 CMR4ST1 CMR4WT1 CMR4PR1 CMR4ST2 CMR4WT2
CMR4PR2 CMR4ST3 CMR4WT3 CMR4PR3 CMR4WT4 CMR4PR4 CMR4WT5
CMR4PR5 CMR4WT6 CMR4PR6 CMR4WT7 CMR4PR7 CMR5ST1 CMR5WT1
CMR5PR1 CMR6ST1 CMR6WT1 CMR6PR1 CMR6ST2 CMR6WT2 CMR6PR2
CMR6ST3 CMR6WT3 CMR6PR3 CMR7ST1 CMR7WT1 CMR7PR1 CMR7ST2
CMR7WT2 CMR7PR2 CMR7ST3 CMR7WT3 CMR7PR3 CMR7WT4 CMR7PR4
CMR8ST1 CMR8WT1 CMR8PR1 CMR8ST2 CMR8WT2 CMR8PR2 CMR8ST3
CMR8WT3 CMR8PR3 CMR9ST1 CMR9WT1 CMR9PR1 CMR9ST2 CMR9WT2
CMR9PR2 CMR0ST1 CMR0WT1 CMR0PR1 CMR0ST2 CMR0WT2 CMR0PR2
CMR0ST3 CMR0WT3 CMR0PR3 CMR0WT4 CMR0PR4 CMR0WT5 CMR0PR5
PNO1ST1 PNO1WT1 PNO1PR1 PNO1ST2 PNO1WT2 PNO1PR2 PNO1ST3
PNO1WT3 PNO1PR3 PNO2ST1 PNO2WT1 PNO2PR1 PNO2ST2 PNO2WT2

PNO2PR2 PNO2ST3 PNO2WT3 PNO2PR3 PNO2WT4 PNO2PR4 PNO2WT5
PNO2PR5 PNO3ST1 PNO3WT1 PNO3PR1 PNO3ST2 PNO3WT2 PNO3PR2
PNO3ST3 PNO3WT3 PNO3PR3 PNO3WT4 PNO3PR4 PNO3WT5 PNO3PR5
PNO4ST1 PNO4WT1 PNO4PR1 PNO4ST2 PNO4WT2 PNO4PR2 PNO4ST3
PNO4WT3 PNO4PR3 PNO5ST1 PNO5WT1 PNO5PR1 PNO5ST2 PNO5WT2
PNO5PR2 PNO5ST3 PNO5WT3 PNO5PR3 PNO6ST1 PNO6WT1 PNO6PR1
PNO6ST2 PNO6WT2 PNO6PR2 PNO6ST3 PNO6WT3 PNO6PR3 PNO7ST1
PNO7WT1 PNO7PR1 PNO7ST2 PNO7WT2 PNO7PR2 PNO7ST3 PNO7WT3
PNO7PR3 PNO7WT4 PNO7PR4 PNO7WT5 PNO7PR5 PDE1ST1 PDE1WT1
PDE1PR1 PDE1ST2 PDE1WT2 PDE1PR2 PDE1ST3 PDE1WT3 PDE1PR3
PDE2ST1 PDE2WT1 PDE2PR1 PDE2ST2 PDE2WT2 PDE2PR2 PDE2ST3
PDE2WT3 PDE2PR3 PDE2WT4 PDE2PR4 PDE3ST1 PDE3WT1 PDE3PR1
PDE3ST2 PDE3WT2 PDE3PR2 PDE3ST3 PDE3WT3 PDE3PR3 PDE3WT4
PDE3PR4 PDE4ST1 PDE4WT1 PDE4PR1 PDE4ST2 PDE4WT2 PDE4PR2
PDE4ST3 PDE4WT3 PDE4PR3 PDE5ST1 PDE5WT1 PDE5PR1 PDE5ST2
PDE5WT2 PDE5PR2 PDE5ST3 PDE5WT3 PDE5PR3 PDE5WT4 PDE5PR4
PDE5WT5 PDE5PR5 PDE6ST1 PDE6WT1 PDE6PR1 PDE6ST2 PDE6WT2
PDE6PR2 PDE6ST3 PDE6WT3 PDE6PR3 PDE7ST1 PDE7WT1 PDE7PR1
PDE7ST2 PDE7WT2 PDE7PR2 PDE7ST3 PDE7WT3 PDE7PR3 PDE7WT4
PDE7PR4 PJA1ST1 PJA1WT1 PJA1PR1 PJA1ST2 PJA1WT2 PJA1PR2
PJA1ST3 PJA1WT3 PJA1PR3 PJA2ST1 PJA2WT1 PJA2PR1 PJA2ST2
PJA2WT2 PJA2PR2 PJA2ST3 PJA2WT3 PJA2PR3 PJA2WT4 PJA2PR4
PJA3ST1 PJA3WT1 PJA3PR1 PJA3ST2 PJA3WT2 PJA3PR2 PJA3ST3
PJA3WT3 PJA3PR3 PJA3WT4 PJA3PR4 PJA4ST1 PJA4WT1 PJA4PR1
PJA4ST2 PJA4WT2 PJA4PR2 PJA4ST3 PJA4WT3 PJA4PR3 PJA4WT4
PJA4PR4 PJA5ST1 PJA5WT1 PJA5PR1 PJA5ST2 PJA5WT2 PJA5PR2
PJA5ST3 PJA5WT3 PJA5PR3 PJA5WT4 PJA5PR4 PJA6ST1 PJA6WT1
PJA6PR1 PJA6ST2 PJA6WT2 PJA6PR2 PJA6ST3 PJA6WT3 PJA6PR3
PJA6WT4 PJA6PR4 PJA6WT5 PJA6PR5 PJA7ST1 PJA7WT1 PJA7PR1
PJA7ST2 PJA7WT2 PJA7PR2 PJA7ST3 PJA7WT3 PJA7PR3 PFE1ST1
PFE1WT1 PFE1PR1 PFE1ST2 PFE1WT2 PFE1PR2 PFE1ST3 PFE1WT3
PFE1PR3 PFE2ST1 PFE2WT1 PFE2PR1 PFE2ST2 PFE2WT2 PFE2PR2
PFE2ST3 PFE2WT3 PFE2PR3 PFE2WT4 PFE2PR4 PFE3ST1 PFE3WT1
PFE3PR1 PFE3ST2 PFE3WT2 PFE3PR2 PFE3ST3 PFE3WT3 PFE3PR3
PFE4ST1 PFE4WT1 PFE4PR1 PFE4ST2 PFE4WT2 PFE4PR2 PFE5ST1
PFE5WT1 PFE5PR1 PFE5ST2 PFE5WT2 PFE5PR2 PFE5ST3 PFE5WT3
PFE5PR3 PFE6ST1 PFE6WT1 PFE6PR1 PFE6ST2 PFE6WT2 PFE6PR2
PFE7ST1 PFE7WT1 PFE7PR1 PFE7ST2 PFE7WT2 PFE7PR2 PFE7ST3
PFE7WT3 PFE7PR3 PMR1ST1 PMR1WT1 PMR1PR1 PMR1ST2 PMR1WT2
PMR1PR2 PMR2ST1 PMR2WT1 PMR2PR1 PMR2ST2 PMR2WT2 PMR2PR2
PMR2ST3 PMR2WT3 PMR2PR3 PMR2WT4 PMR2PR4 PMR2WT5 PMR2PR5
PMR3ST1 PMR3WT1 PMR3PR1 PMR3ST2 PMR3WT2 PMR3PR2 PMR3ST3
PMR3WT3 PMR3PR3 PMR4ST1 PMR4WT1 PMR4PR1 PMR4ST2 PMR4WT2
PMR4PR2 PMR4ST3 PMR4WT3 PMR4PR3 PMR5ST1 PMR5WT1 PMR5PR1
PMR5ST2 PMR5WT2 PMR5PR2 PMR5ST3 PMR5WT3 PMR5PR3 PMR5WT4
PMR5PR4 PMR6ST1 PMR6WT1 PMR6PR1 PMR6ST2 PMR6WT2 PMR6PR2
PMR6ST3 PMR6WT3 PMR6PR3 PMR6WT4 PMR6PR4 PMR6WT5 PMR6PR5
PMR7ST1 PMR7WT1 PMR7PR1 PMR7ST2 PMR7WT2 PMR7PR2 PMR7ST3
PMR7WT3 PMR7PR3 PMR7WT4 PMR7PR4 FNO1ST1 FNO1WT1 FNO1PR1
FNO1ST2 FNO1WT2 FNO1PR2 FNO1ST3 FNO1WT3 FNO1PR3 FNO1WT4
FNO1PR4 FNO1WT5 FNO1PR5 FNO1WT6 FNO1PR6 FNO2ST1 FNO2WT1
FNO2PR1 FNO2ST2 FNO2WT2 FNO2PR2 FNO2ST3 FNO2WT3 FNO2PR3
FNO3ST1 FNO3WT1 FNO3PR1 FNO3ST2 FNO3WT2 FNO3PR2 FDE1ST1
FDE1WT1 FDE1PR1 FDE1ST2 FDE1WT2 FDE1PR2 FDE1ST3 FDE1WT3
FDE1PR3 FDE1WT4 FDE1PR4 FDE1WT5 FDE1PR5 FDE1WT6 FDE1PR6
FDE1WT7 FDE1PR7 FDE1WT8 FDE1PR8 FDE1WT9 FDE1PR9 FDE2ST1

```
FDE2WT1  FDE2PR1  FDE2ST2  FDE2WT2  FDE2PR2  FDE2ST3  FDE2WT3
FDE2PR3  FDE2WT4  FDE2PR4  FDE3ST1  FDE3WT1  FDE3PR1  FDE3ST2
FDE3WT2  FDE3PR2  FDE3ST3  FDE3WT3  FDE3PR3  FJA1ST1  FJA1WT1
FJA1PR1  FJA1ST2  FJA1WT2  FJA1PR2  FJA1ST3  FJA1WT3  FJA1PR3
FJA1WT4  FJA1PR4  FJA1WT5  FJA1PR5  FJA1WT6  FJA1PR6  FJA2ST1
FJA2WT1  FJA2PR1  FJA2ST2  FJA2WT2  FJA2PR2  FJA2ST3  FJA2WT3
FJA2PR3  FJA2WT4  FJA2PR4  FJA3ST1  FJA3WT1  FJA3PR1  FJA3ST2
FJA3WT2  FJA3PR2  FJA3ST3  FJA3WT3  FJA3PR3  FFE1ST1  FFE1WT1
FFE1PR1  FFE1ST2  FFE1WT2  FFE1PR2  FFE1ST3  FFE1WT3  FFE1PR3
FFE1WT4  FFE1PR4  FFE1WT5  FFE1PR5  FFE1WT6  FFE1PR6  FFE1WT7
FFE1PR7  FFE2ST1  FFE2WT1  FFE2PR1  FFE2ST2  FFE2WT2  FFE2PR2
FFE2ST3  FFE2WT3  FFE2PR3  FFE2WT4  FFE2PR4  FFE3ST1  FFE3WT1
FFE3PR1  FFE3ST2  FFE3WT2  FFE3PR2  FMR1ST1  FMR1WT1  FMR1PR1
FMR1ST2  FMR1WT2  FMR1PR2  FMR1ST3  FMR1WT3  FMR1PR3  FMR1WT4
FMR1PR4  FMR1WT5  FMR1PR5  FMR1WT6  FMR1PR6  FMR2ST1  FMR2WT1
FMR2PR1  FMR2ST2  FMR2WT2  FMR2PR2  FMR2ST3  FMR2WT3  FMR2PR3
FMR3ST1  FMR3WT1  FMR3PR1  FMR3ST2  FMR3WT2  FMR3PR2

BEEF1VOL BEEF2VOL BEEF3VOL BEEF4VOL BEEF5VOL BEEF6VOL BEEF7VOL
BEEF8VOL BEEF9VOL POUL1VOL POUL2VOL POUL3VOL POUL4VOL POUL5VOL POUL6VOL
POUL7VOL POUL8VOL POUL9VOL POUL0VOL PORK1VOL PORK2VOL PORK3VOL
PORK4VOL PORK5VOL PORK6VOL PORK7VOL FISH1VOL FISH2VOL FISH3VOL
BEEFVOL  POULVOL  PORKVOL  FISHVOL  BF1VOLAV BF2VOLAV BF3VOLAV
BF4VOLAV BF5VOLAV BF6VOLAV BF7VOLAV BF8VOLAV CH1VOLAV CH2VOLAV
CH3VOLAV CH4VOLAV CH5VOLAV CH6VOLAV CH7VOLAV CH8VOLAV CH9VOLAV
CH0VOLAV PK1VOLAV PK2VOLAV PK3VOLAV PK4VOLAV PK5VOLAV PK6VOLAV
PK7VOLAV FS1VOLAV FS2VOLAV FS3VOLAV BFVOLAV  CHVOLAV  PKVOLAV
FSVOLAV  BF9VOLAV

BEEF1DOL BEEF2DOL BEEF3DOL BEEF4DOL BEEF5DOL BEEF6DOL BEEF7DOL
BEEF8DOL BEEF9DOL POUL1DOL POUL2DOL POUL3DOL POUL4DOL POUL5DOL
POUL6DOL POUL7DOL POUL8DOL POUL9DOL POUL0DOL PORK1DOL PORK2DOL
PORK3DOL PORK4DOL PORK5DOL PORK6DOL PORK7DOL FISH1DOL FISH2DOL
FISH3DOL BEEFDOL  POULDOL  PORKDOL  FISHDOL;

/* this block of code reads in the demographic data. */

infile 'd:\thesis\life.txt' DLM='09'x;

input    FAMNUM    P1Q1    P1Q2    P1Q3    P1Q4    P1Q5
P1Q6     P1Q7      P1Q8    P1Q9    P1Q10   P1Q11
P1Q12    P1Q13     P1Q14   P1Q15   P1Q16   P1Q17
P1Q18    P1Q19     P1Q20   P1Q21   P2LAMB  P2BEEF
P2POUL   P2VEAL    P2FISH  P2PORK  P2OTHER P3BFQ1
P3BFQ2   P3BFQ3    P3BFQ4  P3BFQ5  P3BFQ6  P3BFQ7
P3BFQ8   P3BFQ9    P3BFQ10 P3BFQ11 P3BFQ12
P3BFQ13  P3BFQ14   P3BFQ15 P3BFQ16 P3PKQ1  P3PKQ2
P3PKQ3   P3PKQ4    P3PKQ5  P3PKQ6  P3PKQ7  P3PKQ8
P3PKQ9   P3PKQ10   P3PKQ11 P3PKQ12 P3PKQ13
P3PKQ14  P3PKQ15   P3PKQ16 P3CKQ1  P3CKQ2  P3CKQ3
P3CKQ4   P3CKQ5    P3CKQ6  P3CKQ7  P3CKQ8  P3CKQ9
P3CKQ10  P3CKQ11   P3CKQ12 P3CKQ13 P3CKQ14
P3CKQ15  P3CKQ16   P4Q1    P4Q2    P4Q3    P4Q4
P4Q5     P4Q6      P4Q7    P4Q8    P4Q9    P4Q10
P4Q11    P5Q1      P5Q2    P5Q3    P5Q4    P5Q5
P5Q6     P5Q7      P5Q8    P5Q9    P5Q10   P5Q11
```

```
P5Q12    P5Q13    P5Q14    P5Q15    P5Q16    P5Q17
P5Q18    P5Q19    P5Q20    P5Q21    P5Q22    P5Q23
P5Q24    P5Q25    P5Q26    P5Q27    P5Q28    P5Q29
P5Q30    P5Q31    P5Q32    P5Q33    P5Q34    P5Q35
P5Q36    P5Q37    P5Q38    P5Q39    P5Q40    P5Q41
P5Q42    P5Q43    P5Q44    P5Q45    P5Q46    P5Q47
P5Q48    P5Q49    P5Q50    P5Q51    P5Q52    P5Q53
P5Q54    P5Q55    P5Q56    P5Q57    P5Q58    P5Q59
P5Q60    P5Q61    P5Q62    P5Q63    P5Q64    P5Q65
P6AQ1    P6AQ2    P6AQ3    P6AQ4    P6AQ5    P6AQ6
P6AQ7    P6AQ8    P6AQ9    P6AQ10    P6AQ11    P6AQ12
P6AQ13    P6AQ14    P6AQ15    P6AQ16    P6AQ17    P6B
P6CQ1    P6CQ2    P6CQ3    P6CQ4    P6CQ5    P6CQ6
P6CQ7    P6CQ8    P6DDAY    P6DEVEN    P6DNITE    P6EQ1
P6EQ2    P6EQ3    P6EQ4    P6EQ5    P6EQ6    P6EQ7
P6EQ8    P6EQ9    P6EQ10    P6EQ11    P6EQ12    P6EQ13
P6EQ14    P6EQ15    P6EQ16    P6EQ17    P6EQ18    P6EQ19
P6EQ20    P6EQ21    P6EQ22    P6F    P7Q1    P7Q2
P7Q3    P7Q4    P7Q5    P7Q6    P7Q7    P7Q8
P7Q9    P7Q10    P7Q11    P7Q12    P7Q13    P7Q14
P7Q15    P7Q16    P7Q17    P7Q18    P7Q19    P7Q20
P7Q21    P7Q22    P7Q23    P7Q24    P7Q25    P7Q26
P7Q27    P7Q28    P7Q29    P7Q30    P7Q31    P7Q32
P8Q1    P8Q2    P8Q3    P8Q4    P8Q5    P8Q6
P8Q7    P8Q8    P8Q9    P8Q10    P8Q11    P9AQ1
P9AQ2    P9AQ3    P9AQ4    P9AQ5    P9AQ6    P9AQ7
P9AQ8    P9AQ9    P9AQ10    P9AQ11    P9AQ12    P9AQ13
P9AQ14    P9AQ15    P9BQ1    P9BQ2    P9BQ3    P9BQ4
P9BQ5    P9BQ6    P9BQ7    P9BQ8    P9BQ9    P9BQ10
P9BQ11    P9BQ12    P9BQ13    P9BQ14    P9BQ15    P9CQ1
P9CQ2    P9CQ3    P9CQ4    P9CQ5    P9CQ6    P9CQ7
P9CQ8    P9CQ9    P9CQ10    P9CQ11    P9CQ12    SEX
BICCODE    FEMEDUC    FEMEMP    HHINCOME    ADI
KIDSAGE    MSASIZE    ZIP    CENREG    MALEAGE
MALEDUC    HEADOCC    FEMAGE    HHSIZE    BF1THRDS
BF2THRDS    BF3THRDS    BF4THRDS    BF5THRDS    BF6THRDS
BF7THRDS    BF8THRDS    BF9THRDS    CH1THRDS    CH2THRDS
CH3THRDS    CH4THRDS    CH5THRDS    CH6THRDS    CH7THRDS
CH8THRDS    CH9THRDS    CH0THRDS    PK1THRDS    FS1THRDS
FS2THRDS    FS3THRDS    PK2THRDS    PK3THRDS    PK4THRDS
PK5THRDS    PK6THRDS    PK7THRDS;
run;

/* this block of code manipulates the data to make it ready to use. */

data data1;
set ce2;

/* change the CENREG number depending upon the region. */

*if CENREG<7 or CENREG>7 then delete;
if HHSIZE=0 or HHSIZE=9 then delete;

/* price and expenditure data are normalized by household size. */

B1WT=BEEF1VOL/HHSIZE;
```

```
B2WT=BEEF2VOL/HHSIZE;
B3WT=BEEF3VOL/HHSIZE;
B4WT=BEEF4VOL/HHSIZE;
B5WT=BEEF5VOL/HHSIZE;
B6WT=BEEF6VOL/HHSIZE;
B7WT=BEEF7VOL/HHSIZE;
B8WT=BEEF8VOL/HHSIZE;
B9WT=BEEF9VOL/HHSIZE;
C1WT=POUL1VOL/HHSIZE;
C2WT=POUL2VOL/HHSIZE;
C3WT=POUL3VOL/HHSIZE;
C4WT=POUL4VOL/HHSIZE;
C5WT=POUL5VOL/HHSIZE;
C6WT=POUL6VOL/HHSIZE;
C7WT=POUL7VOL/HHSIZE;
C8WT=POUL8VOL/HHSIZE;
C9WT=POUL9VOL/HHSIZE;
C0WT=POUL0VOL/HHSIZE;
P1WT=PORK1VOL/HHSIZE;
P2WT=PORK2VOL/HHSIZE;
P3WT=PORK3VOL/HHSIZE;
P4WT=PORK4VOL/HHSIZE;
P5WT=PORK5VOL/HHSIZE;
P6WT=PORK6VOL/HHSIZE;
P7WT=PORK7VOL/HHSIZE;
B1DOL=BEEF1DOL/HHSIZE;
B2DOL=BEEF2DOL/HHSIZE;
B3DOL=BEEF3DOL/HHSIZE;
B4DOL=BEEF4DOL/HHSIZE;
B5DOL=BEEF5DOL/HHSIZE;
B6DOL=BEEF6DOL/HHSIZE;
B7DOL=BEEF7DOL/HHSIZE;
B8DOL=BEEF8DOL/HHSIZE;
B9DOL=BEEF9DOL/HHSIZE;
C1DOL=POUL1DOL/HHSIZE;
C2DOL=POUL2DOL/HHSIZE;
C3DOL=POUL3DOL/HHSIZE;
C4DOL=POUL4DOL/HHSIZE;
C5DOL=POUL5DOL/HHSIZE;
C6DOL=POUL6DOL/HHSIZE;
C7DOL=POUL7DOL/HHSIZE;
C8DOL=POUL8DOL/HHSIZE;
C9DOL=POUL9DOL/HHSIZE;
C0DOL=POUL0DOL/HHSIZE;
P1DOL=PORK1DOL/HHSIZE;
P2DOL=PORK2DOL/HHSIZE;
P3DOL=PORK3DOL/HHSIZE;
P4DOL=PORK4DOL/HHSIZE;
P5DOL=PORK5DOL/HHSIZE;
P6DOL=PORK6DOL/HHSIZE;
P7DOL=PORK7DOL/HHSIZE;

/* delete incomplete observations. */

if B1DOL= 0 and B1WT>0 then delete;
if B2DOL= 0 and B2WT>0 then delete;
```

```
if B3DOL= 0 and B3WT>0 then delete;
if B4DOL= 0 and B4WT>0 then delete;
if B5DOL= 0 and B5WT>0 then delete;
if B6DOL= 0 and B6WT>0 then delete;
if B7DOL= 0 and B7WT>0 then delete;
if B8DOL= 0 and B8WT>0 then delete;
if B9DOL= 0 and B9WT>0 then delete;
if C1DOL= 0 and C1WT>0 then delete;
if C2DOL= 0 and C2WT>0 then delete;
if C3DOL= 0 and C3WT>0 then delete;
if C4DOL= 0 and C4WT>0 then delete;
if C5DOL= 0 and C5WT>0 then delete;
if C6DOL= 0 and C6WT>0 then delete;
if C7DOL= 0 and C7WT>0 then delete;
if C8DOL= 0 and C8WT>0 then delete;
if C9DOL= 0 and C9WT>0 then delete;
if C0DOL= 0 and C0WT>0 then delete;
if P1DOL= 0 and P1WT>0 then delete;
if P2DOL= 0 and P2WT>0 then delete;
if P3DOL= 0 and P3WT>0 then delete;
if P4DOL= 0 and P4WT>0 then delete;
if P5DOL= 0 and P5WT>0 then delete;
if P6DOL= 0 and P6WT>0 then delete;
if P7DOL= 0 and P7WT>0 then delete;

/* calculate commodity prices by dividing expenditure by quantity. */

if B1WT=0 then B1PR=0;
else B1PR=B1DOL/B1WT;
if B2WT=0 then B2PR=0;
else B2PR=B2DOL/B2WT;
if B3WT=0 then B3PR=0;
else B3PR=B3DOL/B3WT;
if B4WT=0 then B4PR=0;
else B4PR=B4DOL/B4WT;
if B5WT=0 then B5PR=0;
else B5PR=B5DOL/B5WT;
if B6WT=0 then B6PR=0;
else B6PR=B6DOL/B6WT;
if B7WT=0 then B7PR=0;
else B7PR=B7DOL/B7WT;
if B8WT=0 then B8PR=0;
else B8PR=B8DOL/B8WT;
if B9WT=0 then B9PR=0;
else B9PR=B9DOL/B9WT;
if C1WT=0 then C1PR=0;
else C1PR=C1DOL/C1WT;
if C2WT=0 then C2PR=0;
else C2PR=C2DOL/C2WT;
if C3WT=0 then C3PR=0;
else C3PR=C3DOL/C3WT;
if C4WT=0 then C4PR=0;
else C4PR=C4DOL/C4WT;
if C5WT=0 then C5PR=0;
else C5PR=C5DOL/C5WT;
if C6WT=0 then C6PR=0;
```

```
else C6PR=C6DOL/C6WT;
if C7WT=0 then C7PR=0;
else C7PR=C7DOL/C7WT;
if C8WT=0 then C8PR=0;
else C8PR=C8DOL/C8WT;
if C9WT=0 then C9PR=0;
else C9PR=C9DOL/C9WT;
if C0WT=0 then C0PR=0;
else C0PR=C0DOL/C0WT;
if P1WT=0 then P1PR=0;
else P1PR=P1DOL/P1WT;
if P2WT=0 then P2PR=0;
else P2PR=P2DOL/P2WT;
if P3WT=0 then P3PR=0;
else P3PR=P3DOL/P3WT;
if P4WT=0 then P4PR=0;
else P4PR=P4DOL/P4WT;
if P5WT=0 then P5PR=0;
else P5PR=P5DOL/P5WT;
if P6WT=0 then P6PR=0;
else P6PR=P6DOL/P6WT;
if P7WT=0 then P7PR=0;
else P7PR=P7DOL/P7WT;

/* calculate total expenditures on the beef, poultry, and pork groups. */

BFDOL = B1DOL + B2DOL + B3DOL + B4DOL + B5DOL + B6DOL + B7DOL + B8DOL + B9DOL;
CHDOL = C1DOL + C2DOL + C3DOL + C4DOL + C5DOL + C6DOL + C7DOL + C8DOL + C9DOL + C0DOL;
PKDOL = P1DOL + P2DOL + P3DOL + P4DOL + P5DOL + P6DOL + P7DOL;

/* next line requires that each individual consumes at least one commodity
   each of the groups.  required to make model estimation possible. */

if BFDOL=0 or CHDOL=0 or PKDOL=0 then delete;
run;


/* this block of code determines the sumations and means of the expenditure
   and quantity data, which is used to establish base consumer values and
   in creating the group price indices.  */

proc means sum data=data1;
var B1DOL B2DOL B3DOL B4DOL B5DOL B6DOL B7DOL B8DOL B9DOL
    C1DOL C2DOL C3DOL C4DOL C5DOL C6DOL C7DOL C8DOL C9DOL C0DOL
    P1DOL P2DOL P3DOL P4DOL P5DOL P6DOL P7DOL
    B1WT B2WT B3WT B4WT B5WT B6WT B7WT B8WT B9WT
    C1WT C2WT C3WT C4WT C5WT C6WT C7WT C8WT C9WT C0WT
    P1WT P2WT P3WT P4WT P5WT P6WT P7WT;
output out=data2
sum=SUMB1DOL SUMB2DOL SUMB3DOL SUMB4DOL SUMB5DOL SUMB6DOL SUMB7DOL SUMB8DOL
SUMB9DOL
    SUMC1DOL SUMC2DOL SUMC3DOL SUMC4DOL SUMC5DOL SUMC6DOL SUMC7DOL SUMC8DOL
SUMC9DOL SUMC0DOL
    SUMP1DOL SUMP2DOL SUMP3DOL SUMP4DOL SUMP5DOL SUMP6DOL SUMP7DOL
    SUMB1WT SUMB2WT SUMB3WT SUMB4WT SUMB5WT SUMB6WT SUMB7WT SUMB8WT SUMB9WT
```

```
   SUMC1WT SUMC2WT SUMC3WT SUMC4WT SUMC5WT SUMC6WT SUMC7WT SUMC8WT SUMC9WT
SUMC0WT
   SUMP1WT SUMP2WT SUMP3WT SUMP4WT SUMP5WT SUMP6WT SUMP7WT;

proc means mean data=data1;
var B1DOL B2DOL B3DOL B4DOL B5DOL B6DOL B7DOL B8DOL B9DOL
   C1DOL C2DOL C3DOL C4DOL C5DOL C6DOL C7DOL C8DOL C9DOL C0DOL
   P1DOL P2DOL P3DOL P4DOL P5DOL P6DOL P7DOL
   B1WT B2WT B3WT B4WT B5WT B6WT B7WT B8WT B9WT
   C1WT C2WT C3WT C4WT C5WT C6WT C7WT C8WT C9WT C0WT
   P1WT P2WT P3WT P4WT P5WT P6WT P7WT;
output out=data2b
mean=B1DOL0 B2DOL0 B3DOL0 B4DOL0 B5DOL0 B6DOL0 B7DOL0 B8DOL0 B9DOL0
    C1DOL0 C2DOL0 C3DOL0 C4DOL0 C5DOL0 C6DOL0 C7DOL0 C8DOL0 C9DOL0 C0DOL0
    P1DOL0 P2DOL0 P3DOL0 P4DOL0 P5DOL0 P6DOL0 P7DOL0
    B1WT0 B2WT0 B3WT0 B4WT0 B5WT0 B6WT0 B7WT0 B8WT0 B9WT0
    C1WT0 C2WT0 C3WT0 C4WT0 C5WT0 C6WT0 C7WT0 C8WT0 C9WT0 C0WT0
    P1WT0 P2WT0 P3WT0 P4WT0 P5WT0 P6WT0 P7WT0;

/* this block of code calculates commodity prices for the base consumer. */

data data3;
set data2;

if SUMB1WT=0 then B1PR0=666;
else B1PR0 = SUMB1DOL/SUMB1WT;
if SUMB2WT=0 then B2PR0=666;
else B2PR0 = SUMB2DOL/SUMB2WT;
if SUMB3WT=0 then B3PR0=666;
else B3PR0 = SUMB3DOL/SUMB3WT;
if SUMB4WT=0 then B4PR0=666;
else B4PR0 = SUMB4DOL/SUMB4WT;
if SUMB5WT=0 then B5PR0=666;
else B5PR0 = SUMB5DOL/SUMB5WT;
if SUMB6WT=0 then B6PR0=666;
else B6PR0 = SUMB6DOL/SUMB6WT;
if SUMB7WT=0 then B7PR0=666;
else B7PR0 = SUMB7DOL/SUMB7WT;
if SUMB8WT=0 then B8PR0=666;
else B8PR0 = SUMB8DOL/SUMB8WT;
if SUMB9WT=0 then B9PR0=666;
else B9PR0 = SUMB9DOL/SUMB9WT;
if SUMC1WT=0 then C1PR0=666;
else C1PR0 = SUMC1DOL/SUMC1WT;
if SUMC2WT=0 then C2PR0=666;
else C2PR0 = SUMC2DOL/SUMC2WT;
if SUMC3WT=0 then C3PR0=666;
else C3PR0 = SUMC3DOL/SUMC3WT;
if SUMC4WT=0 then C4PR0=666;
else C4PR0 = SUMC4DOL/SUMC4WT;
if SUMC5WT=0 then C5PR0=666;
else C5PR0 = SUMC5DOL/SUMC5WT;
if SUMC6WT=0 then C6PR0=666;
else C6PR0 = SUMC6DOL/SUMC6WT;
if SUMC7WT=0 then C7PR0=666;
else C7PR0 = SUMC7DOL/SUMC7WT;
```

```
if SUMC8WT=0 then C8PR0=666;
else C8PR0 = SUMC8DOL/SUMC8WT;
if SUMC9WT=0 then C9PR0=666;
else C9PR0 = SUMC9DOL/SUMC9WT;
if SUMC0WT=0 then C0PR0=666;
else C0PR0 = SUMC0DOL/SUMC0WT;
if SUMP1WT=0 then P1PR0=666;
else P1PR0 = SUMP1DOL/SUMP1WT;
if SUMP2WT=0 then P2PR0=666;
else P2PR0 = SUMP2DOL/SUMP2WT;
if SUMP3WT=0 then P3PR0=666;
else P3PR0 = SUMP3DOL/SUMP3WT;
if SUMP4WT=0 then P4PR0=666;
else P4PR0 = SUMP4DOL/SUMP4WT;
if SUMP5WT=0 then P5PR0=666;
else P5PR0 = SUMP5DOL/SUMP5WT;
if SUMP6WT=0 then P6PR0=666;
else P6PR0 = SUMP6DOL/SUMP6WT;
if SUMP7WT=0 then P7PR0=666;
else P7PR0 = SUMP7DOL/SUMP7WT;
run;

/* this block of code creates group prices using the paashe price index. */
/*
data data5;
set data1;
if _n_=1 then
set data3 (keep=B1PR0 B2PR0 B3PR0 B4PR0 B5PR0 B6PR0 B7PR0 B8PR0 B9PR0
           C1PR0 C2PR0 C3PR0 C4PR0 C5PR0 C6PR0 C7PR0 C8PR0 C9PR0 C0PR0
           P1PR0 P2PR0 P3PR0 P4PR0 P5PR0 P6PR0 P7PR0);

BFDOL01 = B1PR0*B1WT + B2PR0*B2WT + B3PR0*B3WT + B4PR0*B4WT + B5PR0*B5WT +
      B6PR0*B6WT + B7PR0*B7WT + B8PR0*B8WT + B9PR0*B9WT;
CHDOL01 = C1PR0*C1WT + C2PR0*C2WT + C3PR0*C3WT + C4PR0*C4WT + C5PR0*C5WT +
      C6PR0*C6WT + C7PR0*C7WT + C8PR0*C8WT + C9PR0*C9WT + C0PR0*C0WT;
PKDOL01 = P1PR0*P1WT + P2PR0*P2WT + P3PR0*P3WT + P4PR0*P4WT + P5PR0*P5WT +
      P6PR0*P6WT + P7PR0*P7WT;
BFDOL11 = B1PR*B1WT + B2PR*B2WT + B3PR*B3WT + B4PR*B4WT + B5PR*B5WT +
      B6PR*B6WT + B7PR*B7WT + B8PR*B8WT + B9PR*B9WT;
CHDOL11 = C1PR*C1WT + C2PR*C2WT + C3PR*C3WT + C4PR*C4WT + C5PR*C5WT +
      C6PR*C6WT + C7PR*C7WT + C8PR*C8WT + C9PR*C9WT + C0PR*C0WT;
PKDOL11 = P1PR*P1WT + P2PR*P2WT + P3PR*P3WT + P4PR*P4WT + P5PR*P5WT +
      P6PR*P6WT + P7PR*P7WT;

PBF = BFDOL11/BFDOL01;
PCH = CHDOL11/CHDOL01;
PPK = PKDOL11/PKDOL01;

if PBF=0 or PCH=0 or PPK=0 then delete;
LNPBF = log(PBF);
LNPCH = log(PCH);
LNPPK = log(PPK);
EXP = BFDOL + CHDOL + PKDOL;
LEXP = log(EXP);

XBF=B1WT+B2WT+B3WT+B4WT+B5WT+B6WT+B7WT+B8WT+B9WT;
```

```
XCH=C1WT+C2WT+C3WT+C4WT+C5WT+C6WT+C7WT+C8WT+C9WT+C0WT;
XPK=P1WT+P2WT+P3WT+P4WT+P5WT+P6WT+P7WT;
run;
*/
/* this block of code creates the tornqvist-theil price index.
   only one of the indices is used at a time.  the other must be
   commented out. */

data data5;
set data1;
if _n_=1 then
set data2b (keep=B1DOL0 B2DOL0 B3DOL0 B4DOL0 B5DOL0 B6DOL0 B7DOL0 B8DOL0 B9DOL0
                C1DOL0 C2DOL0 C3DOL0 C4DOL0 C5DOL0 C6DOL0 C7DOL0 C8DOL0 C9DOL0 C0DOL0
                P1DOL0 P2DOL0 P3DOL0 P4DOL0 P5DOL0 P6DOL0 P7DOL0
                B1WT0 B2WT0 B3WT0 B4WT0 B5WT0 B6WT0 B7WT0 B8WT0 B9WT0
                C1WT0 C2WT0 C3WT0 C4WT0 C5WT0 C6WT0 C7WT0 C8WT0 C9WT0 C0WT0
                P1WT0 P2WT0 P3WT0 P4WT0 P5WT0 P6WT0 P7WT0);
if _n_=1 then
set data3 (keep=B1PR0 B2PR0 B3PR0 B4PR0 B5PR0 B6PR0 B7PR0 B8PR0 B9PR0
               C1PR0 C2PR0 C3PR0 C4PR0 C5PR0 C6PR0 C7PR0 C8PR0 C9PR0 C0PR0
               P1PR0 P2PR0 P3PR0 P4PR0 P5PR0 P6PR0 P7PR0);

BFDOL0 = B1DOL0 + B2DOL0 + B3DOL0 + B4DOL0 + B5DOL0 + B6DOL0 + B7DOL0 + B8DOL0 + B9DOL0;
CHDOL0 = C1DOL0 + C2DOL0 + C3DOL0 + C4DOL0 + C5DOL0 + C6DOL0 + C7DOL0 + C8DOL0 + C9DOL0
       + C0DOL0;
PKDOL0 = P1DOL0 + P2DOL0 + P3DOL0 + P4DOL0 + P5DOL0 + P6DOL0 + P7DOL0;

REG1=0;
If CENREG=1 then REG1=1;
REG2=0;
If CENREG=2 then REG2=1;
REG3=0;
If CENREG=3 then REG3=1;
REG4=0;
If CENREG=4 then REG4=1;
REG5=0;
If CENREG=5 then REG5=1;
REG6=0;
If CENREG=6 then REG6=1;
REG7=0;
If CENREG=7 then REG7=1;
REG8=0;
If CENREG=8 then REG8=1;
REG9=0;
If CENREG=9 then REG9=1;

if B1WT>0 then
PBF1 = (B1PR/B1PR0)**(.5*(B1DOL/BFDOL + B1DOL0/BFDOL0));
else PBF1 = 1;
if B2WT>0 then
PBF2 = (B2PR/B2PR0)**(.5*(B2DOL/BFDOL + B2DOL0/BFDOL0));
else PBF2 = 1;
if B3WT>0 then
PBF3 = (B3PR/B3PR0)**(.5*(B3DOL/BFDOL + B3DOL0/BFDOL0));
else PBF3 = 1;
if B4WT>0 then
```

```
PBF4 = (B4PR/B4PR0)**(.5*(B4DOL/BFDOL + B4DOL0/BFDOL0));
else PBF4 = 1;
if B5WT>0 then
PBF5 = (B5PR/B5PR0)**(.5*(B5DOL/BFDOL + B5DOL0/BFDOL0));
else PBF5 = 1;
if B6WT>0 then
PBF6 = (B6PR/B6PR0)**(.5*(B6DOL/BFDOL + B6DOL0/BFDOL0));
else PBF6 = 1;
if B7WT>0 then
PBF7 = (B7PR/B7PR0)**(.5*(B7DOL/BFDOL + B7DOL0/BFDOL0));
else PBF7 = 1;
if B8WT>0 then
PBF8 = (B8PR/B8PR0)**(.5*(B8DOL/BFDOL + B8DOL0/BFDOL0));
else PBF8 = 1;
if B9WT>0 then
PBF9 = (B9PR/B9PR0)**(.5*(B9DOL/BFDOL + B9DOL0/BFDOL0));
else PBF9 = 1;
if C1WT>0 then
PCH1 = (C1PR/C1PR0)**(.5*(C1DOL/CHDOL + C1DOL0/CHDOL0));
else PCH1 = 1;
if C2WT>0 then
PCH2 = (C2PR/C2PR0)**(.5*(C2DOL/CHDOL + C2DOL0/CHDOL0));
else PCH2 = 1;
if C3WT>0 then
PCH3 = (C3PR/C3PR0)**(.5*(C3DOL/CHDOL + C3DOL0/CHDOL0));
else PCH3 = 1;
if C4WT>0 then
PCH4 = (C4PR/C4PR0)**(.5*(C4DOL/CHDOL + C4DOL0/CHDOL0));
else PCH4 = 1;
if C5WT>0 then
PCH5 = (C5PR/C5PR0)**(.5*(C5DOL/CHDOL + C5DOL0/CHDOL0));
else PCH5 = 1;
if C6WT>0 then
PCH6 = (C6PR/C6PR0)**(.5*(C6DOL/CHDOL + C6DOL0/CHDOL0));
else PCH6 = 1;
if C7WT>0 then
PCH7 = (C7PR/C7PR0)**(.5*(C7DOL/CHDOL + C7DOL0/CHDOL0));
else PCH7 = 1;
if C8WT>0 then
PCH8 = (C8PR/C8PR0)**(.5*(C8DOL/CHDOL + C8DOL0/CHDOL0));
else PCH8 = 1;
if C9WT>0 then
PCH9 = (C9PR/C9PR0)**(.5*(C9DOL/CHDOL + C9DOL0/CHDOL0));
else PCH9 = 1;
if C0WT>0 then
PCH0 = (C0PR/C0PR0)**(.5*(C0DOL/CHDOL + C0DOL0/CHDOL0));
else PCH0 = 1;
if P1WT>0 then
PPK1 = (P1PR/P1PR0)**(.5*(P1DOL/PKDOL + P1DOL0/PKDOL0));
else PPK1 = 1;
if P2WT>0 then
PPK2 = (P2PR/P2PR0)**(.5*(P2DOL/PKDOL + P2DOL0/PKDOL0));
else PPK2 = 1;
if P3WT>0 then
PPK3 = (P3PR/P3PR0)**(.5*(P3DOL/PKDOL + P3DOL0/PKDOL0));
else PPK3 = 1;
```

```
if P4WT>0 then
PPK4 = (P4PR/P4PR0)**(.5*(P4DOL/PKDOL + P4DOL0/PKDOL0));
else PPK4 = 1;
if P5WT>0 then
PPK5 = (P5PR/P5PR0)**(.5*(P5DOL/PKDOL + P5DOL0/PKDOL0));
else PPK5 = 1;
if P6WT>0 then
PPK6 = (P6PR/P6PR0)**(.5*(P6DOL/PKDOL + P6DOL0/PKDOL0));
else PPK6 = 1;
if P7WT>0 then
PPK7 = (P7PR/P7PR0)**(.5*(P7DOL/PKDOL + P7DOL0/PKDOL0));
else PPK7 = 1;

PBF=PBF1*PBF2*PBF3*PBF4*PBF5*PBF6*PBF7*PBF8*PBF9;
PCH=PCH1*PCH2*PCH3*PCH4*PCH5*PCH6*PCH7*PCH8*PCH9*PCH0;
PPK=PPK1*PPK2*PPK3*PPK4*PPK5*PPK6*PPK7;

if PBF=0 or PCH=0 or PPK=0 then delete;
LNPBF = log(PBF);
LNPCH = log(PCH);
LNPPK = log(PPK);
EXP = BFDOL + CHDOL + PKDOL;
LEXP = log(EXP);

XBF=B1WT+B2WT+B3WT+B4WT+B5WT+B6WT+B7WT+B8WT+B9WT;
XCH=C1WT+C2WT+C3WT+C4WT+C5WT+C6WT+C7WT+C8WT+C9WT+C0WT;
XPK=P1WT+P2WT+P3WT+P4WT+P5WT+P6WT+P7WT;

RBF = log(PBF);
RCH = log(PCH);
RPK = log(PPK);
run;

/* this block of code calculates the correlation coefficients.
   this is required to establish if the assumptions of the generalized
   composite commodity theorem are violated. the correlations are computed
   to account for regional differences in the mean. */

data datab1;
set data5;
If B1PR=0 then delete;
RHOBF1=log(B1PR) - RBF;
proc reg data=datab1;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab1a r=NRBF;
proc reg data=datab1;
model RHOBF1=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab1b r=NRHOBF1;
data datab1c;
set datab1a (keep=NRBF rbf);
set datab1b (keep=NRHOBF1 rhobf1);
proc corr data=datab1c cov outp=bf1covar;
var NRBF NRHOBF1;
run;

data datab2;
```

```
set data5;
If B2PR = 0 then delete;
RHOBF2 = log(B2PR) - RBF;
proc reg data=datab2;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab2a r=NRBF;
proc reg data=datab2;
model RHOBF2=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab2b r=NRHOBF2;
data datab2c;
set datab2a (keep=NRBF);
set datab2b (keep=NRHOBF2);
proc corr data=datab2c cov outp=bf2covar;
var NRBF NRHOBF2;
run;

data datab3;
set data5;
If B3PR = 0 then delete;
RHOBF3 = log(B3PR) - RBF;
proc reg data=datab3;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab3a r=NRBF;
proc reg data=datab3;
model RHOBF3=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab3b r=NRHOBF3;
data datab3c;
set datab3a (keep=NRBF);
set datab3b (keep=NRHOBF3);
proc corr data=datab3c cov outp=bf3covar;
var NRBF NRHOBF3;
run;

data datab4;
set data5;
If B4PR = 0 then delete;
RHOBF4 = log(B4PR) - RBF;
proc reg data=datab4;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab4a r=NRBF;
proc reg data=datab4;
model RHOBF4=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab4b r=NRHOBF4;
data datab4c;
set datab4a (keep=NRBF);
set datab4b (keep=NRHOBF4);
proc corr data=datab4c cov outp=bf4covar;
var NRBF NRHOBF4;
run;

data datab5;
set data5;
If B5PR = 0 then delete;
RHOBF5 = log(B5PR) - RBF;
proc reg data=datab5;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
```

```
output out=datab5a r=NRBF;
proc reg data=datab5;
model RHOBF5=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab5b r=NRHOBF5;
data datab5c;
set datab5a (keep=NRBF);
set datab5b (keep=NRHOBF5);
proc corr data=datab5c cov outp=bf5covar;
var NRBF NRHOBF5;
run;

data datab6;
set data5;
If B6PR = 0 then delete;
RHOBF6 = log(B6PR) - RBF;
proc reg data=datab6;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab6a r=NRBF;
proc reg data=datab6;
model RHOBF6=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab6b r=NRHOBF6;
data datab6c;
set datab6a (keep=NRBF);
set datab6b (keep=NRHOBF6);
proc corr data=datab6c cov outp=bf6covar;
var NRBF NRHOBF6;
run;

data datab7;
set data5;
If B7PR = 0 then delete;
RHOBF7 = log(B7PR) - RBF;
proc reg data=datab7;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab7a r=NRBF;
proc reg data=datab7;
model RHOBF7=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab7b r=NRHOBF7;
data datab7c;
set datab7a (keep=NRBF);
set datab7b (keep=NRHOBF7);
proc corr data=datab7c cov outp=bf7covar;
var NRBF NRHOBF7;
run;

data datab8;
set data5;
If B8PR = 0 then delete;
RHOBF8 = log(B8PR) - RBF;
proc reg data=datab8;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab8a r=NRBF;
proc reg data=datab8;
model RHOBF8=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab8b r=NRHOBF8;
data datab8c;
```

```
set datab8a (keep=NRBF);
set datab8b (keep=NRHOBF8);
proc corr data=datab8c cov outp=bf8covar;
var NRBF NRHOBF8;
run;

data datab9;
set data5;
If B9PR = 0 then delete;
RHOBF9 = log(B9PR) - RBF;
proc reg data=datab9;
model RBF=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab9a r=NRBF;
proc reg data=datab9;
model RHOBF9=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datab9b r=NRHOBF9;
data datab9c;
set datab9a (keep=NRBF);
set datab9b (keep=NRHOBF9);
proc corr data=datab9c cov outp=bf9covar;
var NRBF NRHOBF9;
run;

data datac1;
set data5;
if C1PR=0 then delete;
RHOCH1=log(C1PR) - RCH;
proc reg data=datac1;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac1a r=NRCH;
proc reg data=datac1;
model RHOCH1=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac1b r=NRHOCH1;
data datac1c;
set datac1a (keep=NRCH);
set datac1b (keep=NRHOCH1);
proc corr data=datac1c cov outp=ch1covar;
var NRCH NRHOCH1;
run;

data datac2;
set data5;
if C2PR = 0 then delete;
RHOCH2 = log(C2PR) - RCH;
proc reg data=datac2;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac2a r=NRCH;
proc reg data=datac2;
model RHOCH2=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac2b r=NRHOCH2;
data datac2c;
set datac2a (keep=NRCH);
set datac2b (keep=NRHOCH2);
proc corr data=datac2c cov outp=ch2covar;
var NRCH NRHOCH2;
run;
```

```
data datac3;
set data5;
if C3PR = 0 then delete;
RHOCH3 = log(C3PR) - RCH;
proc reg data=datac3;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac3a r=NRCH;
proc reg data=datac3;
model RHOCH3=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac3b r=NRHOCH3;
data datac3c;
set datac3a (keep=NRCH);
set datac3b (keep=NRHOCH3);
proc corr data=datac3c cov outp=ch3covar;
var NRCH NRHOCH3;
run;

data datac4;
set data5;
if C4PR = 0 then delete;
RHOCH4 = log(C4PR) - RCH;
proc reg data=datac4;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac4a r=NRCH;
proc reg data=datac4;
model RHOCH4=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac4b r=NRHOCH4;
data datac4c;
set datac4a (keep=NRCH);
set datac4b (keep=NRHOCH4);
proc corr data=datac4c cov outp=ch4covar;
var NRCH NRHOCH4;
run;

data datac5;
set data5;
if C5PR = 0 then delete;
RHOCH5 = log(C5PR) - RCH;
proc reg data=datac5;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac5a r=NRCH;
proc reg data=datac5;
model RHOCH5=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac5b r=NRHOCH5;
data datac5c;
set datac5a (keep=NRCH);
set datac5b (keep=NRHOCH5);
proc corr data=datac5c cov outp=ch5covar;
var NRCH NRHOCH5;
run;

data datac6;
set data5;
if C6PR = 0 then delete;
RHOCH6 = log(C6PR) - RCH;
```

```
proc reg data=datac6;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac6a r=NRCH;
proc reg data=datac6;
model RHOCH6=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac6b r=NRHOCH6;
data datac6c;
set datac6a (keep=NRCH);
set datac6b (keep=NRHOCH6);
proc corr data=datac6c cov outp=ch6covar;
var NRCH NRHOCH6;
run;

data datac7;
set data5;
if C7PR = 0 then delete;
RHOCH7 = log(C7PR) - RCH;
proc reg data=datac7;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac7a r=NRCH;
proc reg data=datac7;
model RHOCH7=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac7b r=NRHOCH7;
data datac7c;
set datac7a (keep=NRCH);
set datac7b (keep=NRHOCH7);
proc corr data=datac7c cov outp=ch7covar;
var NRCH NRHOCH7;
run;

data datac8;
set data5;
if C8PR = 0 then delete;
RHOCH8 = log(C8PR) - RCH;
proc reg data=datac8;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac8a r=NRCH;
proc reg data=datac8;
model RHOCH8=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac8b r=NRHOCH8;
data datac8c;
set datac8a (keep=NRCH);
set datac8b (keep=NRHOCH8);
proc corr data=datac8c cov outp=ch8covar;
var NRCH NRHOCH8;
run;

data datac9;
set data5;
if C9PR = 0 then delete;
RHOCH9 = log(C9PR) - RCH;
proc reg data=datac9;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac9a r=NRCH;
proc reg data=datac9;
model RHOCH9=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
```

```
output out=datac9b r=NRHOCH9;
data datac9c;
set datac9a (keep=NRCH);
set datac9b (keep=NRHOCH9);
proc corr data=datac9c cov outp=ch9covar;
var NRCH NRHOCH9;
run;

data datac0;
set data5;
if C0PR = 0 then delete;
RHOCH0 = log(C0PR) - RCH;
proc reg data=datac0;
model RCH=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac0a r=NRCH;
proc reg data=datac0;
model RHOCH0=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datac0b r=NRHOCH0;
data datac0c;
set datac0a (keep=NRCH);
set datac0b (keep=NRHOCH0);
proc corr data=datac0c cov outp=ch0covar;
var NRCH NRHOCH0;
run;

data datap1;
set data5;
if P1PR=0 then delete;
RHOPK1=log(P1PR) - RPK;
proc reg data=datap1;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap1a r=NRPK;
proc reg data=datap1;
model RHOPK1=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap1b r=NRHOPK1;
data datap1c;
set datap1a (keep=NRPK);
set datap1b (keep=NRHOPK1);
proc corr data=datap1c cov outp=pk1covar;
var NRPK NRHOPK1;
run;

data datap2;
set data5;
if P2PR = 0 then delete;
RHOPK2 = log(P2PR) - RPK;
proc reg data=datap2;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap2a r=NRPK;
proc reg data=datap2;
model RHOPK2=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap2b r=NRHOPK2;
data datap2c;
set datap2a (keep=NRPK);
set datap2b (keep=NRHOPK2);
proc corr data=datap2c cov outp=pk2covar;
```

```
var NRPK NRHOPK2;
run;

data datap3;
set data5;
if P3PR = 0 then delete;
RHOPK3 = log(P3PR) - RPK;
proc reg data=datap3;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap3a r=NRPK;
proc reg data=datap3;
model RHOPK3=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap3b r=NRHOPK3;
data datap3c;
set datap3a (keep=NRPK);
set datap3b (keep=NRHOPK3);
proc corr data=datap3c cov outp=pk3covar;
var NRPK NRHOPK3;
run;

data datap4;
set data5;
if P4PR = 0 then delete;
RHOPK4 = log(P4PR) - RPK;
proc reg data=datap4;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap4a r=NRPK;
proc reg data=datap4;
model RHOPK4=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap4b r=NRHOPK4;
data datap4c;
set datap4a (keep=NRPK);
set datap4b (keep=NRHOPK4);
proc corr data=datap4c cov outp=pk4covar;
var NRPK NRHOPK4;
run;

data datap5;
set data5;
if P5PR = 0 then delete;
RHOPK5 = log(P5PR) - RPK;
proc reg data=datap5;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap5a r=NRPK;
proc reg data=datap5;
model RHOPK5=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap5b r=NRHOPK5;
data datap5c;
set datap5a (keep=NRPK);
set datap5b (keep=NRHOPK5);
proc corr data=datap5c cov outp=pk5covar;
var NRPK NRHOPK5;
run;

data datap6;
set data5;
```

```
if P6PR = 0 then delete;
RHOPK6 = log(P6PR) - RPK;
proc reg data=datap6;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap6a r=NRPK;
proc reg data=datap6;
model RHOPK6=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap6b r=NRHOPK6;
data datap6c;
set datap6a (keep=NRPK);
set datap6b (keep=NRHOPK6);
proc corr data=datap6c cov outp=pk6covar;
var NRPK NRHOPK6;
run;

data datap7;
set data5;
if P7PR = 0 then delete;
RHOPK7 = log(P7PR) - RPK;
proc reg data=datap7;
model RPK=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap7a r=NRPK;
proc reg data=datap7;
model RHOPK7=REG1 REG2 REG3 REG4 REG5 REG6 REG7 REG8;
output out=datap7b r=NRHOPK7;
data datap7c;
set datap7a (keep=NRPK);
set datap7b (keep=NRHOPK7);
proc corr data=datap7c cov outp=pk7covar;
var NRPK NRHOPK7;
run;

proc rank data=data5 out=data6;
var XBF XCH XPK EXP;
ranks RANKXBF RANKXCH RANKXPK RANKEXP;
run;

data data7;
set data6;
PINTBF = RANKXBF*PBF;
PINTCH = RANKXCH*PCH;
PINTPK = RANKXPK*PPK;
run;

/* this block of code estimates the amount of total expenditures. */

proc reg data=data7;
model LEXP=RANKXBF RANKXCH RANKXPK RANKEXP LNPBF LNPCH LNPPK PINTBF PINTCH PINTPK;
output out=data8 p=LEXPHAT;
run;

/* this block of code establishes normalized group prices and
   budget shares used in estimating the demand model. */

data data9;
SET data8;
```

```
EXPHAT= exp(LEXPHAT);
if EXPHAT < 10 then delete;
P1 = LNPBF - LEXPHAT;
P2 = LNPCH - LEXPHAT;
P3 = LNPPK - LEXPHAT;
M1 = BFDOL/EXP;
M2 = CHDOL/EXP;
M3 = PKDOL/EXP;
run;

/* this block of code specifies the budget share equations from the
   translog model. the equation for M3 is excluded to prevent a singularity. */

proc model data=data9 outparms=data10 maxiter=500 itsur nestit converge=.001;
var M1 M2 P1 P2 P3 ;
parm  A1 .4 A2 .4  B11 B12 B13 B22 B23 B33 ;
M1 = ( A1  + B11*P1 + B12*P2 + B13*P3)/
     ( 1.0 + B11*P1 + B12*P2 + B13*P3 +
             B12*P1 + B22*P2 + B23*P3 +
             B13*P1 + B23*P2 + B33*P3);
M2 = (A2  + B12*P1 + B22*P2 + B23*P3)/
     (1.0 + B11*P1 + B12*P2 + B13*P3 +
            B12*P1 + B22*P2 + B23*P3 +
            B13*P1 + B23*P2 + B33*P3);
fit M1 M2 /outcov outs=data11;

/* this block of code calculates the price and expenditure elasticities. */

NUM1 = A1 + B11*P1 + B12*P2 + B13*P3;
NUM2 = A2 + B12*P1 + B22*P2 + B23*P3;
NUM3 = (1-A1-A2) + B13*P1 + B23*P2 + B33*P3;
DEN =  1 + B11*P1 + B12*P2 + B13*P3
         + B12*P1 + B22*P2 + B23*P3
         + B13*P1 + B23*P2 + B33*P3;

estimate 'EBB' -1 + B11/NUM1 - (B11 + B12 + B13)/DEN;
estimate 'ECC' -1 + B22/NUM2 - (B12 + B22 + B23)/DEN;
estimate 'EPP' -1 + B33/NUM3 - (B13 + B23 + B33)/DEN;
estimate 'EBC' B12/NUM1 - (B12 + B22 + B23)/DEN;
estimate 'EBP' B13/NUM1 - (B13 + B23 + B33)/DEN;
estimate 'ECB' B12/NUM2 - (B11 + B12 + B13)/DEN;
estimate 'ECP' B23/NUM2 - (B13 + B23 + B33)/DEN;
estimate 'EPB' B13/NUM3 - (B11 + B12 + B13)/DEN;
estimate 'EPC' B23/NUM3 - (B12 + B22 + B23)/DEN;
estimate 'EYB' 1 - (B11 + B12 + B13)/NUM1 +
              (B11 + B22 + B33 + 2*B12 + 2*B13 + 2*B23)/DEN;
estimate 'EYC' 1 - (B12 + B22 + B23)/NUM2 +
              (B11 + B22 + B33 + 2*B12 + 2*B13 + 2*B23)/DEN;
estimate 'EYP' 1 - (B13 + B23 + B33)/NUM3 +
              (B11 + B22 + B33 + 2*B12 + 2*B13 + 2*B23)/DEN;
run;
```

VITA

Edward Van Eenoo married his first love, Erica Darnell, May 31, 1997. His second love, Isaac Ray, was born October 14, 1998.